

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
Σχολή Χρηματοοικονομικής και Στατιστικής



Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ ΚΑΙ
ΔΙΟΙΚΗΤΙΚΗ ΚΙΝΔΥΝΟΥ**

**ΠΟΛΥΜΕΤΑΒΛΗΤΗ ΘΕΩΡΙΑ
ΑΚΡΑΙΩΝ ΤΙΜΩΝ ΜΕ ΕΦΑΡΜΟΓΕΣ
ΣΤΗ ΔΙΑΧΕΙΡΙΣΗ ΚΙΝΔΥΝΟΥ**

Απόστολος Κ. Αμέντας

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των α-
παιτήσεων για την απόκτηση του Μεταπτυχιακού Διπλώμα-
τος Ειδίκευσης στην *Αναλογιστική Επιστήμη και Διοικητική
Κινδύνου*

Πειραιάς
Νοέμβριος 2022



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
Σχολή Χρηματοοικονομικής και Στατιστικής



Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ ΚΑΙ
ΔΙΟΙΚΗΤΙΚΗ ΚΙΝΔΥΝΟΥ**

**ΠΟΛΥΜΕΤΑΒΛΗΤΗ ΘΕΩΡΙΑ
ΑΚΡΑΙΩΝ ΤΙΜΩΝ ΜΕ ΕΦΑΡΜΟΓΕΣ
ΣΤΗ ΔΙΑΧΕΙΡΙΣΗ ΚΙΝΔΥΝΟΥ**

Απόστολος Κ. Αμέντας

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των α-
παιτήσεων για την απόκτηση του Μεταπτυχιακού Διπλώμα-
τος Ειδίκευσης στην *Αναλογιστική Επιστήμη και Διοικητική
Κινδύνου*

Πειραιάς
Νοέμβριος 2022

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίστηκε από τη Συνέλευση του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμ. συνεδρίασή της, σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Αναλογιστική Επιστήμη και Διοικητική Κινδύνου.

Τα μέλη της Επιτροπής ήταν:

- Μιχαήλ Μπούτσικας, Αναπληρωτής Καθηγητής (Επιβλέπων)
- Γεώργιος Τζαβελάς, Αναπληρωτής Καθηγητής
- Γεώργιος Πιτσέλης, Επίκουρος Καθηγητής

Η έγκριση της Διπλωματικής Εργασίας από το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς δεν υποδηλώνει αποδοχή των γνώμων του συγγραφέα.

UNIVERSITY OF PIRAEUS
School of Finance and Statistics



Department of Statistics and Insurance Science

**POSTGRADUATE PROGRAM IN
ACTUARIAL SCIENCE AND RISK MANAGEMENT**

**MULTIVARIATE EXTREME
VALUE THEORY WITH
APPLICATIONS TO RISK
MANAGEMENT**

By

Apostolos K. Amentas

MSc Dissertation

submitted to the Department of Statistics and Insurance
Science of the University of Piraeus in partial fulfilment
of the requirements for the degree of Master of Science in
Actuarial Science and Risk Management

Piraeus, Greece
November
2022

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω τον Καθηγητή του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης, κ. Μπούτσικα Μιχαήλ, για τη βοήθεια του στην οργάνωση και την επίβλεψη της παρούσας εργασίας.

ΠΕΡΙΛΗΨΗ

Κύριο αντικείμενο της εργασίας αυτής είναι η παρουσίαση εισαγωγικών αποτελεσμάτων της μονομεταβλητής και πολυμεταβλητής θεωρίας ακραίων τιμών και της σύνδεσής τους μέσω των συναρτήσεων σύζευξης (copulas), καθώς και η εφαρμογή τους στη μέτρηση του κινδύνου σε χαρτοφυλάκια με δύο ή περισσότερες εξαρτημένες επενδύσεις.

Πιο συγκεκριμένα, στο πρώτο κεφάλαιο της παρούσας εργασίας παρατίθενται εισαγωγικές έννοιες και ορισμοί της θεωρίας ακραίων τιμών για τη μονομεταβλητή περίπτωση.

Στο δεύτερο κεφάλαιο παρουσιάζονται αναλυτικοί ορισμοί και ιδιότητες για την εξάρτηση μεταξύ τυχαίων μεταβλητών και για τις συναρτήσεις σύζευξης, καθώς επίσης και για την πολυμεταβλητή θεωρία ακραίων τιμών μέσω των παραμετρικών μεθόδων Μερικών Μεγίστων (Multivariate Block Maxima) και Μέθοδος Υπερβάσεων Κατωφλίου (Multivariate Peaks over Threshold).

Στο τρίτο κεφάλαιο περιγράφονται διάφορα μέτρα κινδύνου, όπως η Μέγιστη Αναμενόμενη Δυνητική Ζημία (Value-at-Risk - VaR) και η Αναμενόμενη Ζημία Ουράς (Expected Shortfall - ES), μέσω διάφορων μεθόδων, όπως η μέθοδος ιστορικής προσομοίωσης, η μέθοδος διακύμανσης-συνδιακύμανσης. Επιπρόσθετα, περιγράφεται η μεθοδολογία εύρεσης του VaR σε επίπεδο χαρτοφυλακίου.

Τέλος, στο τέταρτο κεφάλαιο πραγματοποιείται εφαρμογή της πολυμεταβλητής θεωρίας ακραίων τιμών με πραγματικά δεδομένα δισδιάστατων αποδόσεων μέσω του λογισμικού R. Πιο συγκεκριμένα, οι μέθοδοι εφαρμόζονται σε ένα σύνολο δισδιάστατων αποδόσεων από τα κρυπτονομίσματα Bitcoin (BTC) και Cardano (ADA), με σκοπό την εκτίμηση του VaR μέσω της διμεταβλητής (από κοινού) Συνάρτησης Κατανομής Ακραίων Παρατηρήσεων (Bivariate Extreme Value Distribution) και στη συνέχεια γίνεται σύγκριση με άλλες παραμετρικές μεθόδους εύρεσης της αξίας σε κίνδυνο.

ABSTRACT

The main objective of this MSc thesis is to present introductory results related to univariate and multivariate extreme value theory with the use of copulas, while a second task is to show their application to the measurement of risk in portfolios with two or more dependent investments.

More specifically, the first chapter of this MSc thesis provides introductory concepts and definitions of extreme value theory for the univariate case.

The second chapter presents definitions and properties regarding the dependence between random variables with the use of copulas. It also presents results in multivariate extreme value theory using parametric methods such as *multivariate block maxima* and *multivariate peaks over threshold*.

Chapter three describes various measures of risk, such as Value-at-Risk and Expected Shortfall, and discusses their estimation through various methods, such as historical simulation and variance-covariance method. In addition, a description of the methodology for estimating VaR at the portfolio level is provided.

In chapter four, an application of multivariate extreme value theory is carried out with two-dimensional return data using R software. More specifically, bivariate block maxima methods are applied to a set of two-dimensional returns of Bitcoin (BTC) and Cardano (ADA) cryptocurrencies to estimate VaR via the bivariate (joint) extreme value distribution function. Finally, a numerical comparison study is performed with other parametric methods of Value at Risk estimation.

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1ο: ΜΟΝΟΜΕΤΑΒΛΗΤΗ ΘΕΩΡΙΑ ΑΚΡΑΙΩΝ ΤΙΜΩΝ	1
1.1 ΕΙΣΑΓΩΓΙΚΕΣ ΕΝΝΟΙΕΣ	2
1.1.1 Η Οριακή Κατανομή της Μέγιστης Παρατήρησης	2
1.1.2 Μεγιστοευσταθείς (Max Stable) Κατανομές	2
1.1.3 Κατανομές Ιδίου Τύπου	3
1.1.4 Το Θεμελιώδες Θεώρημα Fisher-Tippett.....	3
1.1.5 Γενικευμένη Κατανομή Ακραιών Τιμών (GEV)	4
1.2 ΜΕΘΟΔΟΣ ΤΟΥ ΜΕΡΙΚΟΥ ΜΕΓΙΣΤΟΥ (BLOCK MAXIMA).....	6
1.2.1 Μοντελοποίηση της Μεθόδου Block Maxima	6
1.2.2 Εκτίμηση των Παραμέτρων της Κατανομής GEV	6
1.2.3 Έλεγχος Καλής Προσαρμογής των Block Maxima στην Κατανομή GEV	7
1.2.4 Εκτίμηση της Αξίας σε Κίνδυνο (VaR) μέσω της Μεθόδου Block Maxima	8
1.3 ΜΕΘΟΔΟΣ PEAKS OVER THRESHOLD (POT).....	9
1.3.1 Μοντελοποίηση της Μεθόδου POT	9
1.3.2 Εκτίμηση των Παραμέτρων της Γενικευμένης Κατανομής Pareto (GPD)	10
1.3.3 Έλεγχος Καλής Προσαρμογής των Block Maxima στην Κατανομή GPD.....	11
1.3.4 Εκτίμηση του VaR μέσω της Μεθόδου POT.....	12
ΚΕΦΑΛΑΙΟ 2ο: ΠΟΛΥΜΕΤΑΒΛΗΤΗ ΘΕΩΡΙΑ ΑΚΡΑΙΩΝ ΤΙΜΩΝ	14
2.1 ΕΞΑΡΤΗΣΗ ΚΑΙ ΣΥΝΑΡΤΗΣΕΙΣ ΣΥΖΕΥΞΗΣ (COPULAS).....	15
2.1.1 Μέτρα Εξάρτησης.....	15
2.1.2 Εξάρτηση μέσω Συναρτήσεων Σύζευξης (Copulas).....	18
2.1.3 Είδη Συναρτήσεων Σύζευξης.....	22
2.2 ΠΟΛΥΜΕΤΑΒΛΗΤΗ ΘΕΩΡΙΑ ΑΚΡΑΙΩΝ ΤΙΜΩΝ.....	26
2.2.1 Πολυμεταβλητή Μέθοδος Block Maxima	27
2.2.2 Πολυμεταβλητή Μέθοδος Peaks over Threshold.....	30
2.2.3 Μοντελοποίηση της Δομής Εξάρτησης	30
ΚΕΦΑΛΑΙΟ 3ο: ΕΠΙΜΕΤΡΗΣΗ ΚΙΝΔΥΝΟΥ	34
3.1 ΜΕΤΡΑ ΚΙΝΔΥΝΟΥ	34
3.2 ΜΕΘΟΔΟΙ ΥΠΟΛΟΓΙΣΜΟΥ ΤΟΥ VaR.....	37
3.3 VaR ΣΕ ΕΠΙΠΕΔΟ ΧΑΡΤΟΦΥΛΑΚΙΟΥ.....	40
ΚΕΦΑΛΑΙΟ 4ο: ΠΡΑΚΤΙΚΗ ΕΦΑΡΜΟΓΗ ΥΠΟΛΟΓΙΣΜΟΥ ΤΟΥ VaR.....	42
4.1 ΕΙΣΑΓΩΓΗ	42
4.2 ΕΚΤΙΜΗΣΗ ΤΟΥ VaR	42
4.3 ΣΥΜΠΕΡΑΣΜΑΤΑ.....	61
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	63

ΚΕΦΑΛΑΙΟ 1ο: ΜΟΝΟΜΕΤΑΒΛΗΤΗ ΘΕΩΡΙΑ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

Η θεωρία ακραίων τιμών χρησιμοποιείται στον κλάδο των χρηματοοικονομικών ως ένα στατιστικό εργαλείο μοντελοποίησης και εκτίμησης μεγάλων ζημιών σε ασφαλιστικά ή επενδυτικά χαρτοφυλάκια.

Η μελέτη των ακραίων παρατηρήσεων σε ένα χαρτοφυλάκιο ζημιών μάς βοηθά να απαντήσουμε σε καθοριστικά ερωτήματα όπως:

- ποια είναι η πιθανότητα η απαίτηση μιας ζημίας, για ένα συγκεκριμένο διάστημα (π.χ. κατά το επόμενο έτος), να ξεπεράσει ένα συγκεκριμένο ποσό.
- τι ύψος ποσού δεν θα ξεπεράσει καμία απαίτηση ζημίας, για ένα συγκεκριμένο διάστημα (π.χ. κατά το επόμενο έτος), με μια συγκεκριμένη πιθανότητα (π.χ. 0,5%).

Για τη μοντελοποίηση της ΘΑΤ θα επικεντρωθούμε σε 2 βασικές παραμετρικές μεθόδους:

- **Μέθοδος Μερικών Μεγίστων (Block Maxima Method)** - χωρίζουμε τα δεδομένα σε ισομεγέθη κομμάτια (blocks) και μελετάμε την οριακή κατανομή των μεγίστων παρατηρήσεων που προκύπτουν από κάθε κομμάτι (block).
- **Μέθοδος Υπερβάσεων Κατωφλίου (Peaks over Threshold Method)** - επιλέγουμε ένα προκαθορισμένο κατώφλι u . Ως μέγιστες παρατηρήσεις θεωρούμε αυτές που υπερβαίνουν το u και μελετάμε την οριακή κατανομή τους.

1.1 ΕΙΣΑΓΩΓΙΚΕΣ ΕΝΝΟΙΕΣ

1.1.1 Η Οριακή Κατανομή της Μέγιστης Παρατήρησης

Πρόταση 1.1: Έστω X_1, X_2, \dots, X_n ανεξάρτητες και ισόνομες τ.μ.(IID) και F η κοινή σ.κ. τους. Αν $M_n = \max\{X_1, X_2, \dots, X_n\}$ τότε η ακριβής σ.κ. της $M_n \forall x \in R$ είναι:

$$P(M_n \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = F(x)^n$$

Παρατήρηση 1.1: Στην πράξη το παραπάνω αποτέλεσμα δεν είναι χρήσιμο, καθώς η F είναι άγνωστη. Πρέπει λοιπόν να προσεγγίσουμε την κατανομή της M_n (όταν $n \rightarrow \infty$) ανεξάρτητα της F .

Πρόταση 1.2: Έστω $x_f = \sup\{x \in R : F(x) < 1\}$ το δεξί άκρο του στηρίγματος της F . Είναι εμφανές ότι $M_n \xrightarrow[n \rightarrow \infty]{} x_f$. Πράγματι:

- Για $x_f < \infty, \forall x \in R \Rightarrow P(M_n \leq x) = F(x)^n \xrightarrow[n \rightarrow \infty]{} \begin{cases} 0, & x < x_f \\ 1, & x \geq x_f \end{cases}$
- Για $x_f = \infty, \forall x \in R \Rightarrow P(M_n \leq x) = F(x)^n \xrightarrow[n \rightarrow \infty]{} 0$

Παρατήρηση 1.2: Το γεγονός ότι η οριακή κατανομή της M_n είναι εκφυλισμένη δεν μας εξυπηρετεί. Ωστόσο, κανονικοποιώντας κατάλληλα το M_n με μεθοδολογία ανάλογη του ΚΟΘ θα μπορούσαμε να οδηγηθούμε στην εύρεση μιας οριακής μη εκφυλισμένης κατανομής για το κανονικοποιημένο μέγιστο. Έστω λοιπόν κατάλληλες ακολουθίες $c_n > 0$ και $b_n \in R$ και μια μη εκφυλισμένη οριακή κατανομή G . Αναζητούμε αποτελέσματα της μορφής:

$$\frac{M_n - b_n}{c_n} \xrightarrow{d} G$$

Τη λύση στην παραπάνω υπόθεση έδωσαν οι Fisher-Tippett (1928) με τη διατύπωση του ομώνυμου θεωρήματος. Ωστόσο, προτού αποτυπώσουμε το θεώρημα, θα δώσουμε μερικούς χρήσιμους για τη συνέχεια ορισμούς.

1.1.2 Μεγιστοευσταθείς (Max Stable) Κατανομές

Ορισμός 1.1: Έστω X_1, X_2, \dots, X_n ανεξάρτητες και ισόνομες τ.μ. (IID) και F η κοινή σ.κ. τους. Η F καλείται **max stable** εάν για κατάλληλη επιλογή ακολουθιών $c_n > 0$ και $b_n \in R$:

$$P\left(\frac{\max\{X_1, X_2, \dots, X_n\} - b_n}{c_n} \leq x\right) = P(X_1 \leq x)$$

Όταν δηλαδή η σ.κ. F των X_i οδηγεί σε κανονικοποιημένο μέγιστο που ακολουθεί την ίδια αρχική κατανομή F .

Πρόταση 1.3: Αν ωστόσο με MD συμβολίσουμε την κλάση όλων των μη-εκφυλισμένων κατανομών που μπορούν να προκύψουν ως όριο κάποιου κανονικοποιημένου μεγίστου, αποδεικνύεται ότι η κλάση κατανομών max stable συμπίπτει με την κλάση κατανομών MD.

$$MD \equiv \text{Max stable}$$

1.1.3 Κατανομές Ιδίου Τύπου

Ορισμός 1.2: Δυο τ.μ. X, Y θα νοείται ότι είναι του ίδιου τύπου αν για κάποια c, b :

$$Y =_d cX + b$$

Παρατήρηση 1.3: Αν υπάρχουν ακολουθίες $c_n > 0$ και $b_n \in R$ και μια μη εκφυλισμένη κατανομή G έτσι ώστε $\frac{M_n - b_n}{c_n} \xrightarrow{d} G$, τότε για διαφορετική κανονικοποίηση, έστω $c'_n = \frac{c_n}{c}$ και $b'_n = b_n - b \frac{c_n}{c}$ με $b \in R$ και $c > 0$, προκύπτει:

$$\frac{M_n - b'_n}{c'_n} = \frac{M_n - b_n + b \frac{c_n}{c}}{\frac{c_n}{c}} = c \frac{M_n - b_n}{c_n} + b \xrightarrow{d} G'$$

Δηλαδή η G' είναι του ίδιου τύπου με την G (βλ. ορισμό 1.2)

1.1.4 Το Θεμελιώδες Θεώρημα Fisher-Tippett

Θεώρημα 1.1: (Fisher-Tippett, 1928) Έστω X_1, X_2, \dots, X_n μια ακολουθία από IID τ.μ. και $M_n = \max\{X_1, X_2, \dots, X_n\}$. Αν υπάρχουν ακολουθίες $c_n > 0$ και $b_n \in R$ και μια μη εκφυλισμένη κατανομή G ώστε:

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n - b_n}{c_n} \leq x\right) = G(x), x \in R$$

τότε η G είναι ίδιου τύπου με μια από τις παρακάτω 3 κατανομές:

$$\text{Gumbel: } \Lambda(x) = e^{-e^{-x}}, x \in R$$

$$\text{Frechet: } \Phi_a(x) = \begin{cases} 0 & , x \leq 0 \\ e^{-x^{-a}} & , x > 0 \end{cases} , a > 0$$

$$\text{Weibull: } \Psi_a(x) = \begin{cases} e^{-(-x)^a} & , x \leq 0 \\ 1 & , x > 0 \end{cases} , a > 0$$

Οι Λ, Φ_a, Ψ_a καλούνται τυπικές κατανομές ακροτάτων και οι ιδίου τύπου με αυτές καλούνται κατανομές ακροτάτων.

Παρατήρηση 1.4: Αποδεικνύεται ότι οι τρεις παραπάνω τύποι κατανομών είναι οι μοναδικοί που ικανοποιούν τον ορισμό των max stable κατανομών (βλ. ορισμό 1.1). Ως εκ τούτου, σε κάθε κατανομή των X_i θα αντιστοιχεί μια μοναδική οριακή κατανομή (αρκεί το κανονικοποιημένο μέγιστο $\frac{M_n - b_n}{c_n}$ να συγκλίνει) εκ των παραπάνω τριών. Η διαδικασία εύρεσης των κατανομών που αντιστοιχούν σε μια εκ των Λ, Φ_a, Ψ_a απαιτεί περαιτέρω ανάλυση και δεν μας απασχολεί ιδιαίτερα αφού στην πράξη ακολουθείται μια ασφαλέστερη μεθοδολογία για την εύρεση της οριακής κατανομής.

1.1.5 Γενικευμένη Κατανομή Ακραίων Τιμών (GEV)

Ορισμός 1.3: Η παραμετρική οικογένεια κατανομών που ενοποιεί τις τρεις οικογένειες κατανομών ακροτάτων von Mises (1936), Jenkinson (1955) Λ, Φ_a, Ψ_a (βλ. θεώρημα 1.1) καλείται κατανομή *GEV* με παραμέτρους μ, σ, ξ όπου $\mu, \xi \in R$ και $\sigma > 0$ και είναι της μορφής:

$$G(x) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]_+^{-\frac{1}{\xi}} \right\}$$

όπου:

- μ , η παράμετρος θέσης
- σ , η παράμετρος κλίμακας
- ξ , η παράμετρος σχήματος

Πρόταση 1.4: Για κατάλληλες τιμές των παραμέτρων μ, σ, ξ η κατανομή *GEV* μετασχηματίζεται σε μια εκ των τριών κατανομών ακροτάτων Λ, Φ_a, Ψ_a

- για $\xi \rightarrow 0, \mu = b, \sigma = c$

$$G(x) \xrightarrow{\xi \rightarrow 0} e^{-e^{-\frac{x-b}{c}}} = \Lambda \left(\frac{x-b}{c} \right) , \quad x \in R$$

- για $\xi = \frac{1}{\alpha}$ ($\xi > 0$), $\mu = c + b$, $\sigma = \frac{c}{\alpha}$

$$G(x) = e^{-\left(\frac{x-b}{c}\right)^{-\alpha}} = \Phi_\alpha\left(\frac{x-b}{c}\right), \quad \frac{x-b}{c} > 0 \Leftrightarrow 1 + \xi \frac{x-\mu}{\sigma} > 0$$

- για $\xi = -\frac{1}{\alpha}$ ($\xi < 0$), $\mu = -c + b$, $\sigma = \frac{c}{\alpha}$

$$G(x) = e^{-\left(\frac{x-b}{c}\right)^{\alpha}} = \Psi_\alpha\left(\frac{x-b}{c}\right), \quad \frac{x-b}{c} < 0 \Leftrightarrow 1 + \xi \frac{x-\mu}{\sigma} > 0$$

Θεώρημα 1.2: Έστω X_1, X_2, \dots, X_n μια ακολουθία από IID τ.μ. και $M_n = \max\{X_1, X_2, \dots, X_n\}$. Αν υπάρχουν ακολουθίες $c_n > 0$ και $b_n \in R$ και μια μη εκφυλισμένη κατανομή G ώστε:

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n - b_n}{c_n} \leq x\right) = G(x)$$

τότε η $G \sim GEV$ με παραμέτρους μ, σ, ξ όπου $\mu, \xi \in R$ και $\sigma > 0$.

Πρόταση 1.5: Αν ισχύει $P\left(\frac{M_n - b_n}{c_n} \leq x\right) \xrightarrow{n \rightarrow \infty} G \sim GEV(x; \mu, \sigma, \xi)$ (βλ. Θεώρημα 1.2), τότε και το μη-κανονικοποιημένο μέγιστο ισχύει (προσεγγιστικά, για μεγάλο n) ότι $M_n \sim GEV(y; \mu_n, \sigma_n, \xi)$.

Πράγματι αν θέσουμε $y = c_n x + b_n$

$$\begin{aligned} P\left(\frac{M_n - b_n}{c_n} \leq x\right) &= P(M_n \leq y) \approx GEV\left(\frac{y - b_n}{c_n}; \mu, \sigma, \xi\right) \\ &= e^{-\left(1 + \xi \frac{y - b_n - \mu}{\sigma}\right)^{\frac{-1}{\xi}}} = e^{-\left(1 + \xi \frac{y - (b_n + \mu c_n)}{\sigma c_n}\right)^{\frac{-1}{\xi}}} \end{aligned}$$

αν τώρα θέσουμε $\mu_n = b_n + \mu c_n$ και $\sigma_n = \sigma c_n$ τότε:

$$GEV\left(\frac{y - b_n}{c_n}; \mu, \sigma, \xi\right) = e^{-\left(1 + \xi \frac{y - \mu_n}{\sigma_n}\right)^{\frac{-1}{\xi}}} = GEV(y; \mu_n, \sigma_n, \xi)$$

και επομένως:

$$P(M_n \leq y) \approx GEV(y; \mu_n, \sigma_n, \xi)$$

Παρατήρηση 1.5: Μέσω της GEV δεν χρειάζεται να επιλέξουμε εκ των προτέρων μία από τις τρεις κατανομές ακροτάτων και μετά να εκτιμήσουμε τις παραμέτρους. Η εκτίμηση των παραμέτρων γίνεται απευθείας από την GEV , με την οριακή κατανομή ακροτάτων να προκύπτει από την τιμή της παραμέτρου σχήματος ξ (βλ. Πρόταση 1.4).

1.2 ΜΕΘΟΔΟΣ ΤΟΥ ΜΕΡΙΚΟΥ ΜΕΓΙΣΤΟΥ (BLOCK MAXIMA)

1.2.1 Μοντελοποίηση της Μεθόδου Block Maxima

Ορισμός 1.4: Έστω X_1, X_2, \dots, X_m μια ακολουθία από IID τ.μ. που προέρχονται από μια άγνωστη κατανομή F . Αν χωρίσουμε τις m το πλήθος παρατηρήσεις σε k το πλήθος υποσύνολα που το καθένα περιέχει n το πλήθος παρατηρήσεις ($m = kn$), τότε οι μέγιστες παρατηρήσεις από κάθε υποσύνολο (k το πλήθος) ονομάζονται μερικά μέγιστα (block maxima – B.M.) και είναι της μορφής:

$$Y_i = M_n^{(i)} = \max(X_1^{(i)}, X_2^{(i)}, \dots, X_n^{(i)}), \quad i = 1, 2, \dots, k$$

Πόρισμα 1.1: Από το Θεώρημα 1.2, την Πρόταση 1.5 καθώς και τον Ορισμό 1.4 προκύπτει ότι για μεγάλο n (παρατηρήσεις σε κάθε block):

$$Y_i \sim GEV(y; \mu_n, \sigma_n, \xi)$$

1.2.2 Εκτίμηση των Παραμέτρων της Κατανομής GEV

Παρατήρηση 1.6: Στην παρούσα εργασία θα επικεντρωθούμε στη μέθοδο της μέγιστης πιθανοφάνειας. Στα θετικά της μεθόδου είναι ότι χρησιμοποιείται ευρέως και προσαρμόζεται εύκολα στην κατασκευή σύνθετων μοντέλων ενώ στα αρνητικά της μεθόδου είναι ότι οι ασυμπτωτικές ιδιότητες των εκτιμητών μέγιστης πιθανοφάνειας (ε.μ.π.) δεν ισχύουν σε όλο τον παραμετρικό χώρο της GEV . Το πρόβλημα αυτό μελετήθηκε από τον Smith (1985) ο οποίος κατέληξε ότι:

- όταν $\xi > -0.5$, οι ε. μ. π. έχουν τις ασυμπτωτικές τους ιδιότητες.
- όταν $-1 < \xi < -0.5$, οι ε. μ. π. δεν έχουν τις ασυμπτωτικές τους ιδιότητες.
- όταν $\xi < -1$, είναι πιθανό οι ε.μ.π. να μην μπορούν να βρεθούν

Στην πράξη αυτός ο περιορισμός δεν αποτελεί πρόβλημα, αφού η ΘΑΤ στον κλάδο των χρηματοοικονομικών ασχολείται κυρίως με κατανομές που έχουν βαριές ή λεπτές ουρές που εκτείνονται στο άπειρο.

Πρόταση 1.6: Έστω $Y_1, Y_2, \dots, Y_k \sim GEV(y; \mu, \sigma, \xi)$. Η συνάρτηση του λογαρίθμου της πιθανοφάνειας είναι:

$$\ell(\mu, \sigma, \xi) = \ln \prod_{i=1}^k g(y_i; \mu, \sigma, \xi) = \sum_{i=1}^k \ln g(y_i; \mu, \sigma, \xi)$$

επειδή:

$$g(y; \mu, \sigma, \xi) = \frac{d}{dy} GEV(y; \mu, \sigma, \xi) = \frac{1}{\sigma} \left(1 + \xi \frac{y - \mu}{\sigma}\right)^{\frac{-1}{\xi} - 1} e^{-(1 + \xi \frac{y - \mu}{\sigma})^{\frac{-1}{\xi}}}$$

δοθέντος ότι $1 + \xi \frac{y - \mu}{\sigma} > 0$, τότε για $\xi \neq 0$

$$l(\mu, \sigma, \xi) = -k \ln \sigma - \left(\frac{1}{\xi} + 1\right) \sum_{i=1}^k \ln \left(1 + \xi \frac{y_i - \mu}{\sigma}\right) - \sum_{i=1}^k \left(1 + \xi \frac{y_i - \mu}{\sigma}\right)^{\frac{-1}{\xi}}$$

δοθέντος ότι $1 + \xi \frac{y - \mu}{\sigma} > 0$.

Ενώ για $\xi = 0$

$$l(\mu, \sigma, 0) = -k \ln \sigma - \sum_{i=1}^k \frac{y_i - \mu}{\sigma} - \sum_{i=1}^k e^{-\left(\frac{y_i - \mu}{\sigma}\right)}$$

Παρατήρηση 1.7: Η λύση του παραπάνω συστήματος είναι αδύνατη με αναλυτικές εκφράσεις είτε λόγω περιορισμών (όταν $\xi \neq 0 \Rightarrow 1 + \xi \frac{y - \mu}{\sigma} > 0$) είτε λόγω πεπλεγμένης μορφής (όταν $\xi = 0$). Ωστόσο είναι δυνατό να πάρουμε τιμές για τους ε.μ.π. μέσω αριθμητικών μεθόδων βελτιστοποίησης.

Παρατήρηση 1.8: Η εύρεση δ.ε. για τους ε.μ.π. είναι εξαιρετικά σημαντική. Ιδιαίτερα για τον ε.μ.π ξ είναι απαραίτητη. Ο ακριβής προσδιορισμός της κατανομής ακροτάτων εξαρτάται από την παράμετρο ξ και μικρές αποκλίσεις στην τιμή του ξ μπορεί να οδηγήσουν σε μη ασφαλή συμπεράσματα. Για περισσότερες λεπτομέρειες παραπέμπουμε στο βιβλίο του Coles S. (2001) καθώς και στις σημειώσεις του Μπούτσικα Μ.(2008).

1.2.3 Έλεγχος Καλής Προσαρμογής των Block Maxima στην Κατανομή GEV

Όπως αναφέραμε παραπάνω, για μεγάλο n (θεωρητικά $n \rightarrow \infty$) τα B.M. ακολουθούν ασυμπτωτικά την κατανομή *GEV*. Ωστόσο σε πραγματικές εφαρμογές το n είναι πεπερασμένο. Ως εκ τούτου πρέπει να ελεγχθεί (για δοθέν n) αν τα B.M. ακολουθούν την κατανομή *GEV* με τις εκτιμηθείσες παραμέτρους.

Πρόταση 1.7: Το Quantile – Quantile (Q-Q) διάγραμμα των B.M. ως προς την κατανομή

$GEV(\hat{\mu}, \hat{\sigma}, \hat{\xi})$ αποτελείται από τα σημεία:

$$\left(G^{-1}\left(\frac{i}{k+1}; \hat{\mu}, \hat{\sigma}, \hat{\xi}\right), Y_i \right), i = 1, \dots, k$$

όπου, $Y_{(1)} < Y_{(2)} < \dots < Y_{(k)}$ είναι τα B.M. διατεταγμένα από το μικρότερο στο μεγαλύτερο. Αν τα παραπάνω σημεία βρίσκονται κοντά στη διαγώνιο του γραφήματος $\left(Y_i \approx G^{-1}\left(\frac{i}{k+1}, \hat{\mu}, \hat{\sigma}, \hat{\xi}\right) \right)$ τότε μπορούμε να θεωρήσουμε ότι έχουμε καλή προσαρμογή των B.M. στην κατανομή GEV . ($B.M. \sim GEV$)

Παρατήρηση 1.9: Προσοχή χρειάζεται κατά την επιλογή των k το πλήθος κομματιών (blocks). Από τη σχέση $m = kn$ προκύπτει ότι αν το k είναι πολύ μεγάλο υπάρχει κίνδυνος το n να είναι μικρό και να μην ισχύει $Y_i \sim GEV$. Πρέπει λοιπόν να επιλέγεται το μέγιστο δυνατό k που θα δίνει ένα ικανοποιητικό εύρος B.M. και παράλληλα θα εξασφαλίζει ότι τα B.M. προσαρμόζονται ικανοποιητικά στην GEV . Για περισσότερες λεπτομέρειες παραπέμπουμε στο βιβλίο του Coles S. (2001) καθώς και στις σημειώσεις του Μπούτσικα Μ. (2008).

1.2.4 Εκτίμηση της Αξίας σε Κίνδυνο (VaR) μέσω της Μεθόδου Block Maxima

Ορισμός 1.5: Το VaR μιας ζημίας X με επίπεδο σημαντικότητας $p \in (0,1)$ είναι το κατώφλι που ξεπερνάει η ζημία X με πιθανότητα p :

$$P(X > VaR_p(X)) = p \Leftrightarrow VaR_p(X) = F_X^{-1}(1 - p)$$

Πρόταση 1.8: Έστω X_i μια ζημία όπου $i = 1, 2, \dots, n$ είναι η χρονική διάρκεια (κάθε πότε εμφανίζεται η ζημία, π.χ. ημέρα). Αν $M_n = \max\{X_1, X_2, \dots, X_n\}$ τότε για το $VaR_p(X)$ ισχύει ότι:

$$F_{M_n}(VaR_p(X)) = (1 - p)^n$$

Πράγματι:

$$F_{M_n}(VaR_p(X)) = P(M_n \leq VaR_p(X)) = P(\max\{X_1, X_2, \dots, X_n\} \leq VaR_p(X))$$

και επειδή X_1, X_2, \dots, X_n είναι IID τ.μ.:

$$F_{M_n}(VaR_p(X)) = P(X \leq VaR_p(X))^n = F_X(VaR_p(X))^n$$

όμως δείξαμε ότι $VaR_p(X) = F_X^{-1}(1 - p)$ και άρα:

$$F_{M_n}(VaR_p(X)) = F_X(F_X^{-1}(1-p))^n = (1-p)^n$$

Πόρισμα 1.2: Από την Πρόταση 1.8 και την Πρόταση 1.5 προκύπτει ότι για μεγάλο n η $F_{M_n} \approx GEV$ και συνεπώς:

$$F_{M_n}(VaR_p(X)) \approx GEV(VaR_p(X)) = (1-p)^n$$

$$VaR_p(X) \approx GEV^{-1}(1-p)^n$$

Πρόταση 1.9: Η εκτίμηση για το VaR προκύπτει ως το άνω p ποσοστημόριο της προσαρμοσμένης στα δεδομένα κατανομής GEV με παραμέτρους τους ε.μ.π. $\hat{\mu}, \hat{\sigma}, \hat{\xi}$ και είναι της μορφής:

$$\widehat{VaR}_p(X) = GEV^{-1}((1-p)^n; \hat{\mu}, \hat{\sigma}, \hat{\xi})$$

και συνεπώς:

$$\widehat{VaR}_p(X) = \begin{cases} \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}(1 - (-\ln(1-p)^n)^{-\hat{\xi}}), & \xi \neq 0 \\ \hat{\mu} - \hat{\sigma} \ln(-\ln(1-p)^n) & , \quad \xi = 0 \end{cases}$$

1.3 ΜΕΘΟΔΟΣ PEAKS OVER THRESHOLD (POT)

1.3.1 Μοντελοποίηση της Μεθόδου POT

Ορισμός 1.6: Έστω X_1, X_2, \dots, X_n μια ακολουθία από IID τ.μ. και $M_n = \max\{X_1, X_2, \dots, X_n\}$. Η σ.κ. F των X_i ανήκει στην περιοχή έλξης της κατανομής GEV (De Haan et al, 2006) και θα γράφουμε $F \in MDA(\xi)$ αν υπάρχουν ακολουθίες $c_n > 0$ και $b_n \in R$ τέτοιες ώστε:

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n - b_n}{c_n} \leq x\right) = G(x), x \in R$$

Ορισμός 1.7: Έστω X_1, X_2, \dots, X_n μια ακολουθία από IID τ.μ. όπου F είναι η κοινή σ.κ. τους και u ένα κατώφλι. Η τ.μ. $X_i - u \mid X_i > u$, όπου $i = 1, 2, \dots, j$ καλείται υπέρβαση της τ.μ. X_i ενώ η σ.κ. της καλείται κατανομή υπέρβασης και είναι της μορφής:

$$F_u(z) = P(X - u \leq z \mid X > u) = \frac{F(u+z) - F(u)}{1 - F(u)}, \quad z > 0$$

Παρατήρηση 1.10: Στην πράξη το παραπάνω αποτέλεσμα δεν είναι χρήσιμο καθώς η F_u βασίζεται στην F , η οποία όμως είναι άγνωστη. Πρέπει λοιπόν να βασιστούμε στην οριακή κατανομή της F_u .

Ορισμός 1.8: Ορίζουμε ως Γενικευμένη κατανομή Pareto τη σ.κ.:

$$GPD(z; \sigma, \xi) = \begin{cases} 1 - \left(1 + \xi \frac{z}{\sigma}\right)^{-\frac{1}{\xi}}, & \xi \neq 0 \\ 1 - e^{-\frac{z}{\sigma}}, & \xi = 0 \end{cases}$$

όπου $\sigma > 0$ είναι η παράμετρος κλίμακας και ξ η παράμετρος σχήματος. Το στήριγμα της παραπάνω κατανομής αποτελείται (αναλόγως της τιμής της παραμέτρου ξ) από όλα τα $z \in R$, για τα οποία ισχύει:

$$\begin{cases} 0 < z < -\frac{\sigma}{\xi}, & \xi < 0 \\ z > 0, & \xi \geq 0 \end{cases}$$

Θεώρημα 1.3: Έστω X_1, X_2, \dots, X_n μια ακολουθία από IID τ.μ. όπου F είναι η κοινή σ.κ. τους και ισχύει ότι $F \in MDA(\xi)$, τότε για μεγάλο u , θα ισχύει ότι:

$$F_u(z) = P(X - u \leq z | X > u) \approx GPD(z; \sigma_u, \xi)$$

όπου $\sigma_u = \sigma + \xi(u - \mu)$ η παράμετρος κλίμακας, η οποία αποτελεί συνάρτηση των παραμέτρων της GEV και ξ η παράμετρος σχήματος, η οποία είναι ίδια και στις δύο κατανομές.

Παρατήρηση 1.11: Μια δυσκολία της μεθόδου POT είναι η επιλογή που πρέπει να κάνουμε για το κατώφλι u . Η βέλτιστη επιλογή του u είναι καθοριστικής σημασίας καθώς αν πάρουμε u πολύ μικρό υπάρχει ο κίνδυνος η $F_u \approx GPD$, ενώ αν πάρουμε u υπερβολικά μεγάλο οι υπερβάσεις θα είναι πολύ λίγες και ως εκ τούτου υπάρχει κίνδυνος να μην μπορούμε να κάνουμε ασφαλείς εκτιμήσεις. Για περισσότερες πληροφορίες σχετικά με τις μεθόδους εύρεσης του βέλτιστου u , παραπέμπουμε στον Coles S. (2001).

1.3.2 Εκτίμηση των Παραμέτρων της Γενικευμένης Κατανομής Pareto (GPD)

Όπως και στην περίπτωση της GEV , οι παράμετροι της GPD θα βρεθούν με τη μέθοδο της μέγιστης πιθανοφάνειας. Για την GPD θα συμβολίσουμε τις υπερβάσεις των X_i , όπου X_i είναι μεγαλύτερα από το u , με Y_1, Y_2, \dots, Y_k , δηλαδή:

$$\{Y_1, Y_2, \dots, Y_k\} = \{X_i - u | X_i > u, i = 1, 2, \dots, n\}$$

Παρατήρηση 1.10: Έστω $Y_1, Y_2, \dots, Y_k \sim F_u \approx GPD(y; \sigma_u, \xi)$. Η συνάρτηση του λογαρίθμου της πιθανοφάνειας είναι:

$$l(\sigma_u, \xi) = \ln \prod_{i=1}^k h(y_i; \sigma_u, \xi) = \sum_{i=1}^k \ln h(y_i; \sigma_u, \xi)$$

επειδή:

$$h(y; \sigma_u, \xi) = \frac{d}{dy} GPD(y; \mu, \sigma_u, \xi) = \frac{1}{\sigma_u} \left(1 + \xi \frac{y}{\sigma_u}\right)^{-\frac{1}{\xi}-1}$$

δοθέντος ότι $1 + \xi \frac{y}{\sigma_u} > 0$, τότε για $\xi \neq 0$:

$$l(\sigma_u, \xi) = -k \ln \sigma_u - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^k \ln \left(1 + \xi \frac{y_i}{\sigma_u}\right)$$

δοθέντος ότι $1 + \xi \frac{y}{\sigma_u} > 0$. Ενώ για $\xi = 0$:

$$l(\sigma_u, 0) = -k \ln \sigma_u - \frac{1}{\sigma_u} \sum_{i=1}^k y_i$$

Παρατήρηση 1.12: Όταν το $\xi = 0$ η συνάρτηση του λογαρίθμου της πιθανοφάνειας μεγιστοποιείται στο $\widehat{\sigma}_u = \frac{\sum_{i=1}^k y_i}{k}$ ενώ όταν το $\xi \neq 0$ δεν υπάρχουν αναλυτικές εκφράσεις για τους ε.μ.π. των σ_u, ξ . Ωστόσο μπορούν να βρεθούν με τη χρήση αριθμητικών μεθόδων βελτιστοποίησης.

1.3.3 Έλεγχος Καλής Προσαρμογής των Block Maxima στην Κατανομή GPD

Όπως αιτιολογήσαμε παραπάνω, για μεγάλο u , οι υπερβάσεις των X_i πάνω από το u ακολουθούν προσεγγιστικά την κατανομή GPD . Ωστόσο για δοθέν u πρέπει να ελεγχθεί αν τα POT προσαρμόζονται ικανοποιητικά στην κατανομή GPD με τις εκτιμηθείσες παραμέτρους. Αυτό μπορεί να γίνει με τη χρήση διαγραμμάτων Q-Q plot.

Πρόταση 1.11: Το Q-Q διάγραμμα των POT ως προς την κατανομή $GPD(\widehat{\sigma}_u, \widehat{\xi})$ αποτελείται από τα σημεία:

$$\left(H^{-1}\left(\frac{i}{k+1}; \widehat{\sigma}_u, \widehat{\xi}\right), Y_i\right), \quad i = 1, \dots, k$$

όπου, $Y_{(1)} < Y_{(2)} < \dots < Y_{(k)}$ είναι τα POT διατεταγμένα από το μικρότερο στο

μεγαλύτερο. Αν τα παραπάνω σημεία βρίσκονται κοντά στη διαγώνιο του γραφήματος τότε μπορούμε να θεωρήσουμε πως έχουμε καλή προσαρμογή των POT στην κατανομή *GPD*.

1.3.4 Εκτίμηση του VaR μέσω της Μεθόδου POT

Παραπάνω ορίσαμε ότι το *VaR* μιας ζημίας X με επίπεδο σημαντικότητας $p \in (0,1)$ είναι το κατώφλι που ξεπερνάει η ζημία X με πιθανότητα p , δηλαδή:

$$P(X > VaR_p(X)) = p \Leftrightarrow VaR_p(X) = F_X^{-1}(1 - p)$$

Μπορούμε λοιπόν μέσω της μεθόδου POT να εκτιμήσουμε την πιθανότητα να ξεπεράσει μια παρατήρηση το *VaR* κάθε $1/p$ χρονικές περιόδους.

Πρόταση 1.12: Αν $\{Y_1, Y_2, \dots, Y_k\} = \{X_i - u \mid X_i > u \ i = 1, 2, \dots, n\}$ είναι οι υπερβάσεις των X_i πάνω από ένα κατώφλι u αρκετά μεγάλο τότε δείξαμε ότι $Y_1, Y_2, \dots, Y_k \sim F_u \approx GPD(y; \sigma_u, \xi)$ δηλαδή ισχύει ότι:

$$P(X - u > y \mid X > u) \approx 1 - GPD(y; \sigma_u, \xi)$$

συνεπώς ισχύει και ότι:

$$P(X > y \mid X > u) \approx 1 - GPD(y - u; \sigma_u, \xi)$$

Αν τώρα θέσουμε όπου $y = VaR_p(X)$ προκύπτει ότι:

$$P(X > VaR_p(X) \mid X > u) \approx 1 - GPD(VaR_p(X) - u; \sigma_u, \xi) \Rightarrow$$

$$P(X > VaR_p(X)) \approx P(X > u) \left(1 + \xi \frac{VaR_p(X) - u}{\sigma_u} \right)^{-\frac{1}{\xi}}$$

όμως γνωρίζουμε ότι $P(X > VaR_p(X)) = p$ και ότι $P(X > u) \approx \frac{k}{n}$ (είναι ο λόγος των παρατηρήσεων πάνω από το κατώφλι u προς το σύνολο των παρατηρήσεων) και άρα λύνοντας ως προς $VaR_p(X)$ προκύπτει ότι:

$$VaR_p(X) \approx u + \frac{\sigma_u}{\xi} \left(\left(\frac{k}{np} \right)^\xi - 1 \right)$$

Πρόταση 1.13: Θεωρώντας ότι το p είναι αρκετά μικρό ώστε $VaR_p(X) > u$, η εκτίμηση για το VaR θα είναι της μορφής:

$$\widehat{VaR}_p(X) = \begin{cases} u + \frac{\widehat{\sigma}_u}{\widehat{\xi}} \left(\left(\frac{k}{np} \right)^{\widehat{\xi}} - 1 \right), & \widehat{\xi} \neq 0 \\ u + \widehat{\sigma}_u \ln \left(\frac{k}{np} \right), & \widehat{\xi} = 0 \end{cases}$$

ΚΕΦΑΛΑΙΟ 2ο:

ΠΟΛΥΜΕΤΑΒΛΗΤΗ ΘΕΩΡΙΑ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

Τα ακραία γεγονότα είναι πολύ συχνά σε πολλούς τομείς της αναλογιστικής και χρηματοοικονομικής επιστήμης. Ζημίες σε ασφαλιστικά χαρτοφυλάκια περιουσίας ή υγείας μπορεί περιστασιακά να είναι τόσο μεγάλες ώστε να έχουν σημαντικές οικονομικές επιπτώσεις στον ασφαλιστή ή στον αντασφαλιστή. Αντίστοιχα, ακραίες διακυμάνσεις στα επιτόκια ή στις τιμές των μετοχών ή κρυπτονομισμάτων μπορεί επίσης να επηρεάσουν δραματικά τις αξίες των περιουσιακών στοιχείων ή/και των υποχρεώσεων ενός χρηματοπιστωτικού ιδρύματος.

Σε πολλές περιπτώσεις, τα ακραία γεγονότα συμβαίνουν σε περισσότερες από μία τυχαίες μεταβλητές. Για παράδειγμα, μια ασφαλιστική εταιρεία μπορεί να ενδιαφέρεται για απώλειες που συμβαίνουν σε διαφορετικές περιφέρειες, ένας επενδυτής μπορεί να ενδιαφέρεται για τη διακύμανση ενός χαρτοφυλακίου με διάφορες μετοχές, ενώ μια ασφαλιστική εταιρεία ζωής μπορεί να ενδιαφέρεται για τη θνησιμότητα πολλαπλών ζώων που ασφαλίζονται στο ίδιο συμβόλαιο. Ενώ είναι βολικό να υποθέσουμε ότι οι σχετικές τυχαίες μεταβλητές είναι ανεξάρτητες, τις περισσότερες φορές η θεώρηση αυτή είναι ακατάλληλη.

Συνεπώς, είναι απαραίτητο να μοντελοποιηθεί η δομή εξάρτησης αυτών των τυχαίων μεταβλητών, οι οριακές τους κατανομές και τελικά η από κοινού συνάρτηση κατανομής τους. Μια προφανής επιλογή θα ήταν η θεώρηση ότι οι τυχαίες μεταβλητές που εξετάζουμε ακολουθούν μια από κοινού πολυμεταβλητή κανονική κατανομή. Ωστόσο, στην περίπτωση μελέτης ακραίων γεγονότων, δεν υπάρχει κάποιος προφανής λογικός τρόπος να ορισθεί μια δομή εξάρτησης μεταξύ δύο η περισσότερων οριακών κατανομών.

Οι συναρτήσεις σύζευξης (copulas) έχουν γίνει πολύ δημοφιλείς ως εργαλείο για την αναπαράσταση της δομής εξάρτησης τυχαίων μεταβλητών ξεχωριστά από τις οριακές κατανομές τους.

2.1 ΕΞΑΡΤΗΣΗ ΚΑΙ ΣΥΝΑΡΤΗΣΕΙΣ ΣΥΖΕΥΞΗΣ (COPULAS)

2.1.1 Μέτρα Εξάρτησης

Ενώ η έννοια της εξάρτησης των τυχαίων μεταβλητών είναι διαισθητικά σαφής, είναι λιγότερο σαφές πώς μπορεί να ποσοτικοποιηθεί. Σε αυτήν την ενότητα θα παραθέσουμε μερικά από τα πιο γνωστά μέτρα εξάρτησης. Έπειτα, θα παρουσιάσουμε τη συνάρτηση σύζευξης, η οποία περιγράφει πλήρως τη δομή εξάρτησης που σχετίζεται με την από κοινού συνάρτηση κατανομής δύο ή περισσότερων τυχαίων μεταβλητών.

Ορισμός 2.1: Έστω X_1, X_2 δυο τυχαίες μεταβλητές με F την από κοινού συνάρτηση κατανομής τους, F_1, F_2 οι περιθώριες συναρτήσεις κατανομών και θετικές πεπερασμένες διακυμάνσεις. Η γραμμική εξάρτηση ή δείκτης συσχέτισης Pearson χρησιμοποιείται ως μέτρο εξάρτησης δύο τυχαίων μεταβλητών και είναι της μορφής:

$$\rho(X_1, X_2) = \frac{\mathbb{E}[(X_1 - \mathbb{E}(X_1))(X_2 - \mathbb{E}(X_2))]}{\sqrt{\text{Var}(X_1)}\sqrt{\text{Var}(X_2)}}$$

όπου $-1 \leq \rho(X_1, X_2) \leq 1$. Η τέλεια θετική συσχέτιση επιτυγχάνεται από την τιμή 1 ενώ η τέλεια αρνητική από την τιμή -1 .

Παρατήρηση 2.1: Βασικό μειονέκτημα του δείκτη Pearson είναι το ότι επικεντρώνεται στη μελέτη της γραμμικής εξάρτησης. Παραμένει αμετάβλητος μόνο κάτω από έναν αυστηρά αυξανόμενο γραμμικό μετασχηματισμό των μεταβλητών, όμως αυτό δεν ισχύει απαραίτητα για έναν μη γραμμικό μετασχηματισμό.

Γενικότερα, σε χρηματοοικονομικές μεταβλητές και ειδικά όταν αυτές παίρνουν ακραίες τιμές συναντάμε μια απρόβλεπτη εξάρτηση, η οποία ξεφεύγει κατά πολύ από τη γραμμική εξάρτηση. Ένα καλό μέτρο εξάρτησης μεταξύ δύο μεταβλητών θα πρέπει να έχει την ιδιότητα να παραμένει αμετάβλητο όταν εφαρμόζουμε την ίδια μονότονη συνάρτηση σε κάθε μεταβλητή. Συνεπώς ο δείκτης Pearson δεν είναι ένα κατάλληλο μέτρο εξάρτησης τυχαίων μεταβλητών με ακραίες παρατηρήσεις. Εναλλακτικά, η μέτρηση της εξάρτησης δυο μεταβλητών μπορεί να γίνει με βάση τη θέση των παραγόντων των μεταβλητών όταν αυτοί διατάσσονται σε σειρά.

Ορισμός 2.2: Έστω $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ ένα δείγμα n παρατηρήσεων επί του διανύσματος (X, Y) των τυχαίων μεταβλητών X, Y . Αν:

$$R(x_i): \text{ο βαθμός ή τάξη (rank) της } x_i, \quad i = 1, \dots, n$$

$R(y_i)$: ο βαθμός ή τάξη (rank) της y_i , $i = 1, \dots, n$

τότε ο συντελεστής συσχέτισης του Spearman δίνεται από τον τύπο:

$$\rho_s(X, Y) = \frac{\sum_{i=1}^n [R(x_i) - \bar{R}_x][R(y_i) - \bar{R}_y]}{\sqrt{\sum_{i=1}^n [R(x_i) - \bar{R}_x]^2} \sqrt{\sum_{i=1}^n [R(y_i) - \bar{R}_y]^2}}$$

όπου $-1 \leq \rho_s(X, Y) \leq 1$. Η τέλεια θετική συσχέτιση επιτυγχάνεται από την τιμή 1 ενώ η τέλεια αρνητική από την τιμή -1 . Υπενθυμίζεται ότι με $R(x_i)$ συμβολίζεται η θέση της παρατήρησης x_i στο διατεταγμένο δείγμα των x_1, \dots, x_n (π.χ. αν οι παρατηρήσεις (x_i, y_i) , $i = 1, 2, 3$ είναι οι $(2, 6), (4, 5), (3, 4)$ τότε τα αντίστοιχα $(R(x_i), R(y_i))$ είναι $(1, 3), (3, 2), (2, 1)$) ενώ με \bar{R}_x συμβολίζεται το:

$$\frac{R(x_1) + \dots + R(x_n)}{n} = \frac{1 + \dots + n}{n} = \frac{n+1}{2}.$$

Παρατήρηση 2.2: Ο δείκτης του Spearman είναι επί της ουσίας ο δείκτης του Pearson υπολογιζόμενος όμως με βάση τις τάξεις των παρατηρήσεων και όχι αυτές καθαυτές τις παρατηρήσεις. Παραμένει αμετάβλητος ακόμη και κάτω από έναν αυστηρά αυξανόμενο μη γραμμικό μετασχηματισμό των μεταβλητών αφού ακόμη και αν η τιμή μιας παρατήρησης αλλάξει, η σχετική κατάταξή της παραμένει η ίδια.

Τα παραπάνω μέτρα εξάρτησης παρέχουν πληροφορία για τη συνολική εξάρτηση δυο τυχαίων μεταβλητών. Πολλές φορές ωστόσο, μας ενδιαφέρει περισσότερο μια στοχευμένη μελέτη της εξάρτησης στις ουρές των κατανομών. Για παράδειγμα, μια χρηματιστηριακή κρίση μπορεί να οδηγήσει σε ακραίες ζημίες σε μια σειρά επενδύσεων του ίδιου επενδυτικού χαρτοφυλακίου. Ως εκ τούτου, η δομή εξάρτησης δύο η περισσότερων μεταβλητών στα άκρα των κατανομών έχει βαρύνουσα σημασία και μπορεί να μετρηθεί από τους συντελεστές εξάρτησης άνω και κάτω ουράς.

Ορισμός 2.3: Ορίζουμε τον συντελεστή εξάρτησης άνω ουράς ως:

$$\lambda_U = \lim_{u \rightarrow 1^-} P(X > F_X^{-1}(u) | Y > F_Y^{-1}(u))$$

όπου:

- λ_U είναι μια οριακή πιθανότητα, οπότε παίρνει μια τιμή μεταξύ 0 και 1.
- Ο λ_U δείχνει εάν οι υψηλές τιμές μιας τυχαίας μεταβλητής X , τείνουν να συνδέονται με υψηλές τιμές μιας άλλης τυχαίας μεταβλητής Y .

- Ο λ_U δίνει την πιθανότητα η τυχαία μεταβλητή X να λάβει μια τιμή στην άνω ουρά της κατανομής (π.χ. δεξιά ουρά με μάζα πιθανότητας 5% που σημαίνει ότι $u = 0.95$), δεδομένου ότι η τυχαία μεταβλητή Y λαμβάνει μια τιμή στο ίδιο επίπεδο της επάνω ουράς της κατανομής της.

Για τον συντελεστή λ_u ισχύει ότι αν ισούται με μηδέν τότε οι μεταβλητές X και Y είναι ανεξάρτητες στην άνω ουρά. Αντίστοιχα, αν $\lambda_u \in (0,1]$ τότε οι μεταβλητές X και Y είναι εξαρτημένες στην άνω ουρά.

Ορισμός 2.4: Ορίζουμε τον συντελεστή εξάρτησης κάτω ουράς ως:

$$\lambda_L = \lim_{u \rightarrow 0^+} P(X < F_X^{-1}(u) | Y \leq F_Y^{-1}(u))$$

όπου:

- λ_L είναι μια οριακή πιθανότητα και η τιμή του κυμαίνεται μεταξύ 0 και 1.
- Ο λ_L δείχνει εάν οι χαμηλές τιμές της μίας τυχαίας μεταβλητής X τείνουν να συνδέονται με χαμηλές τιμές της άλλης τυχαίας μεταβλητής Y .
- Ο λ_L εξετάζει την πιθανότητα της τυχαίας μεταβλητής X να λαμβάνει μια τιμή από την κάτω ουρά της κατανομής της (π.χ. μια αριστερή ουρά με μάζα πιθανότητας 5% που σημαίνει ότι $u = 0.05$), δεδομένου ότι η τυχαία μεταβλητή Y λαμβάνει μια τιμή στο ίδιο επίπεδο της κάτω ουράς της αντίστοιχης κατανομής της.

Ορισμός 2.5: Το τ του Kendall για το τυχαίο διάνυσμα (X, Y) ορίζεται ως εξής:

$$\tau(X, Y) = P\{(X - \tilde{X})(Y - \tilde{Y}) > 0\} - P\{(X - \tilde{X})(Y - \tilde{Y}) < 0\}$$

όπου $(X, Y), (\tilde{X}, \tilde{Y})$ είναι ανεξάρτητα διανύσματα από την κατανομή του τ.δ. (X, Y) .

Από τον τύπο του Kendall διαπιστώνουμε ότι ουσιαστικά το τ αποτελεί την πιθανότητα της συμφωνίας μείον την πιθανότητα της ασυμφωνίας.

Θεώρημα 2.1: Έστω (X, Y) διάνυσμα τυχαίων μεταβλητών με συνάρτηση σύζευξης (copula) C , τότε το τ του Kendall για το διάνυσμα δίνεται από τον παρακάτω τύπο:

$$\tau(X, Y) = Q(C, C) = 4 \iint_{[0,1]^2} C(u, v) dC(u, v) - 1$$

Το παραπάνω ολοκλήρωμα είναι η μέση τιμή της τυχαίας μεταβλητής $C(U, V)$ όπου $U, V \sim U(0,1)$.

$$\tau(X, Y) = 4E(C(U, V)) - 1$$

2.1.2 Εξάρτηση μέσω Συναρτήσεων Σύζευξης (Copulas)

Η από κοινού συνάρτηση κατανομής συνδυάζει τις πληροφορίες από τις οριακές κατανομές και τον τρόπο με τον οποίο οι μεταβλητές εξαρτώνται η μία από την άλλη. Ωστόσο, δεν εκφράζει αυτή την εξάρτηση εμφανώς. Δεν μπορούμε να διακρίνουμε αμέσως τη φύση της αλληλεξάρτησης παρατηρώντας απλώς τον τύπο της από κοινού συνάρτησης κατανομής. Οι συναρτήσεις σύζευξης (copulas) παρέχουν μια εναλλακτική προσέγγιση που εκφράζει ρητά την αλληλεξάρτηση μεταξύ των μεταβλητών. Μας επιτρέπουν να αποδημήσουμε την από κοινού κατανομή ενός συνόλου μεταβλητών σε συνιστώσες (τις περιθώριες κατανομές συν μια copula) που μπορούν να προσαρμοστούν μεμονωμένα. Η συνάρτηση σύζευξης (copula) λοιπόν, αποτελεί μια συνάρτηση, η οποία εκφράζει μια πολυμεταβλητή αθροιστική συνάρτηση κατανομής μέσω των επιμέρους περιθώριων συναρτήσεων κατανομών κάθε τυχαίας μεταβλητής. Στην ουσία η συνάρτηση σύζευξης (copula) εισάγει ως δεδομένα περιθώριες πιθανότητες τυχαίων μεταβλητών και παράγει μια από κοινού συνάρτηση κατανομής.

Ορισμός 2.6: Για μια διδιάστατη κατανομή η συνάρτηση σύζευξης (copula) είναι μια συνάρτηση C_{XY} και ορίζεται ως:

$$C_{XY}[F_X(x), F_Y(y)] = P(X \leq x, Y \leq y) = F_{X,Y}(x, y)$$

Πιο συνοπτικά ο παραπάνω τύπος γράφεται ως εξής:

$$C[u, v] = F_{X,Y}(x, y) \text{ όπου } u = F_X(x) \text{ και } v = F_Y(y)$$

Για πολυδιάστατη ανάλυση, δηλαδή για την περίπτωση όπου εξετάζουμε παραπάνω από δύο τυχαίες μεταβλητές, η συνάρτηση λαμβάνει την παρακάτω μορφή:

$$C[u_1, u_2, \dots, u_d] = F_{X_1, X_2, \dots, X_d}(x_1, x_2, \dots, x_d) \text{ όπου } u_i = F_{X_i}(x_i)$$

Τα μεγέθη u_1, u_2, \dots, u_d καθώς και η συνάρτηση C λαμβάνουν τιμές στο διάστημα $[0,1]$, διότι αποτελούν τιμές πιθανοτήτων.

Πρόταση 2.1: Οι συναρτήσεις σύζευξης (copulas) διαθέτουν τρεις τεχνικές ιδιότητες, οι οποίες διασφαλίζουν την ορθή περιγραφή μιας από κοινού συνάρτησης κατανομής τυχαίων μεταβλητών.

1. Η συνάρτηση σύζευξης (copula) αποτελεί μια αυξανόμενη συνάρτηση των δεδομένων της:

$$C[u_1, \dots, u_i^*, \dots, u_d] > C[u_1, \dots, u_i, \dots, u_d] \text{ όπου } u_i^* > u_i \text{ και } i = 1, \dots, d$$

Η παραπάνω ιδιότητα δικαιολογείται από πιθανοθεωρητικής πλευράς, διότι εάν $u_i^* > u_i$ τότε $P(X_i \leq x_i^*) > P(X_i \leq x_i)$ για τα αντίστοιχα $x_i^* = F_{X_i}^{-1}(u_i^*)$ και $x_i = F_{X_i}^{-1}(u_i)$, άρα $x_i^* > x_i$ και επομένως:

$$P(X_1 \leq x_1, \dots, X_i \leq x_i^*, \dots, X_d \leq x_d) > P(X_1 \leq x_1, \dots, X_i \leq x_i, \dots, X_d \leq x_d)$$

2. Αν όλες οι περιθώριες συναρτήσεις κατανομών λαμβάνουν την τιμή 1 εκτός από μία, τότε η συνάρτηση C λαμβάνει την τιμή αυτής της συνάρτησης κατανομής:

$$C(1, \dots, 1, u_i, 1, \dots, 1) = u_i \text{ για } i = 1, 2, \dots, d \text{ και } u_i \in [0, 1]$$

Η παραπάνω ιδιότητα δικαιολογείται, διότι $u_k = 1 \Rightarrow P(X_k \leq x_k) = 1$ (δηλ. απόλυτη βεβαιότητα), για $x_k = F_{X_k}^{-1}(u_k)$ και η μοναδική αβεβαιότητα στην ανωτέρω από κοινού πιθανότητα είναι η τιμή της περιθώριας συνάρτησης κατανομής της i -οστής τυχαίας μεταβλητής.

3. Μια συνάρτηση σύζευξης (copula) έχει πάντα μία «έγκυρη» τιμή πιθανότητας, δηλαδή:

$$C(u_1, u_2, \dots, u_d) \in [0, 1]$$

Ένα σημαντικό θεώρημα στη θεωρία των συναρτήσεων σύζευξης (copulas) αποτελεί το θεώρημα του Sklar (1959), στο οποίο αναφέρεται ότι η κατασκευή της εξάρτησης μεταξύ τυχαίων μεταβλητών, μπορεί να πραγματοποιηθεί μέσω των συναρτήσεων σύζευξης (copulas).

Θεώρημα 2.1: Έστω F μία από κοινού συνάρτηση κατανομής, αποτελούμενη από περιθώριες συναρτήσεις κατανομών F_1, \dots, F_d . Υπάρχει μία συνάρτηση σύζευξης (copula) C , τέτοια ώστε, για όλα τα $x_1, \dots, x_d \in [-\infty, \infty]$:

$$F(x_1, \dots, x_d) = C[F_1(x_1), \dots, F_d(x_d)]$$

Στην περίπτωση που οι μεταβλητές ακολουθούν συνεχή κατανομή, τότε η συνάρτηση σύζευξης (copula) είναι μοναδική.

Παρατήρηση 2.3: Το θεώρημα του Sklar μας πληροφορεί ότι αν διαθέτουμε μια από κοινού συνάρτηση κατανομής και τις αντίστοιχες περιθώριες συναρτήσεις κατανομής, τότε υπάρχει μια συνάρτηση σύζευξης (copula) και μέσω αυτής έχουμε τη δυνατότητα να περιγράψουμε την εξάρτηση και τη συμπεριφορά των τυχαίων αυτών μεταβλητών.

Πρόταση 2.2: Το αντίθετο του θεωρήματος 2.1 επίσης ισχύει, δηλαδή αν C είναι μια συνάρτηση σύζευξης (copula) και F_1, \dots, F_d είναι οι αθροιστικές συναρτήσεις κατανομής κάθε μεταβλητής, τότε η συνάρτηση F που ορίστηκε παραπάνω είναι μια από κοινού

αθροιστική συνάρτηση κατανομής με περιθώριες αθροιστικές συναρτήσεις κατανομής τις F_1, \dots, F_d .

Όπως αναφέραμε προηγουμένως, οι συναρτήσεις σύζευξης (copulas) μπορούν να χρησιμοποιηθούν για να ορίσουμε την εξάρτηση μεταξύ των μεταβλητών σε συνδυασμό με τους συντελεστές που αναφέραμε στο κεφάλαιο 2.1.1.

Παρατήρηση 2.4: Υπενθυμίζουμε τον συντελεστή εξάρτησης κάτω ουράς λ_L ο οποίος ισούται με:

$$\lambda_L = \lim_{u \rightarrow 0^+} P(X < F_X^{-1}(u) | Y \leq F_Y^{-1}(u))$$

και μετράει το όριο αυτής της πιθανότητας όσο το u τείνει στο μηδέν και άνω. Ο συντελεστής μπορεί να εκφραστεί με διαφορετικό τρόπο και ως εξής:

$$\lambda_L = \lim_{u \rightarrow 0^+} P(X < x | Y \leq y)$$

όπου x και y είναι τα ποσοστά κάτω ουράς της κατανομής των τυχαίων μεταβλητών X και Y αντίστοιχα με:

$$P(X \leq x_u) = P(Y \leq y_u) = u$$

Με τη χρήση του ορισμού των εξαρτημένων πιθανοτήτων ο συντελεστής κάτω ουράς ισούται τελικά με:

$$\lambda_L = \lim_{u \rightarrow 0^+} \frac{P(X \leq x_u, Y \leq y_u)}{P(Y \leq y_u)}$$

Ο αριθμητής ισούται με τη συνάρτηση σύζευξης (copula) $C[u, u]$, διότι αποτελεί μία από κοινού αθροιστική συνάρτηση κατανομής πιθανότητας, ενώ ο παρονομαστής εκφράζει την πιθανότητα το Y να λάβει τιμές μικρότερες του y_u , το οποίο γνωρίζουμε ότι είναι το u .

Ορισμός 2.7: Επομένως, ο συντελεστής εξάρτησης κάτω ουράς λαμβάνει την παρακάτω απλούστερη και πιο εύχρηστη μορφή:

$$\lambda_L = \lim_{u \rightarrow 0^+} \frac{C[u, u]}{u}$$

Με αυτόν τον τρόπο παρατηρούμε ότι η εξάρτηση των τυχαίων μεταβλητών μπορεί να υπολογιστεί άμεσα με τη χρήση των συναρτήσεων σύζευξης (copulas).

Ορισμός 2.8: Ορίζουμε τη λεγόμενη συνάρτηση σύζευξης επιβίωσης, η οποία ορίζεται ως εξής:

$$\bar{F}(x, y) = P(X > x, Y > y) = \bar{C}[\bar{F}_X(x), \bar{F}_Y(y)]$$

όπου $\bar{F}_X(x) = 1 - F_X(x)$ και $\bar{F}_Y(y) = 1 - F_Y(y)$.

Ορισμός 2.9: Με τη χρήση της σύζευξης επιβίωσης μπορούμε να ορίσουμε και τον συντελεστή εξάρτησης άνω ουράς λ_U ως:

$$\lambda_U = \lim_{u \rightarrow 1^-} \frac{\bar{C}[1 - u, 1 - u]}{1 - u} = \lim_{u \rightarrow 0^+} \frac{\bar{C}[u, u]}{u}$$

Παρατήρηση 2.4: Ισχύει ότι:

$$P(X \leq x \text{ ή } Y \leq y) = P(X \leq x) + P(Y \leq y) - P(X \leq x, Y \leq y)$$

$$1 - P(X > x, Y > y) = P(X \leq x) + P(Y \leq y) - P(X \leq x, Y \leq y)$$

$$P(X > x, Y > y) = 1 - P(X \leq x) - P(Y \leq y) + P(X \leq x, Y \leq y)$$

Επομένως, η συνάρτηση σύζευξης (copula) επιβίωσης σχετίζεται με την αρχική συνάρτηση copula ως εξής:

$$\bar{C}[1 - u, 1 - v] = 1 - u - v + C[u, v]$$

Με λίγη περισσότερη ανάλυση για τον συντελεστή εξάρτησης άνω ουράς μπορούμε να τον εκφράσουμε, σύμφωνα και με την αρχική συνάρτηση σύζευξης (copula) C . Αρχικά γνωρίζουμε ότι ο λ_U ισούται με:

$$\lambda_U = \lim_{u \rightarrow 1^-} P(X > x_u | Y > y_u)$$

όπου τα x_u, y_u αποτελούν τα άνω όρια των συναρτήσεων κατανομής των τυχαίων μεταβλητών X, Y αντίστοιχα και γνωρίζουμε ότι:

$$P(X \leq x_u) = P(Y \leq y_u) = u$$

και:

$$P(X > x_u) = P(Y > y_u) = 1 - u$$

Επομένως ο συντελεστής λ_U προκύπτει ως:

$$\lambda_U = \lim_{u \rightarrow 1^-} \frac{P(X > x_u, Y > y_u)}{P(Y > y_u)}$$

Γνωρίζουμε ότι ο αριθμητής ισούται με $1 - 2u + C[u, u]$, ενώ ο παρονομαστής ισούται με την πιθανότητα το Y να υπερβεί την τιμή y_u , το οποίο επίσης γνωρίζουμε ότι ισούται με $1 - u$. Τελικά, για τον συντελεστή εξάρτησης άνω ουράς καταλήγουμε στην εξής εξίσωση:

$$\lambda_U = \lim_{u \rightarrow 1^-} \frac{1 - 2u + C[u, u]}{1 - u}$$

2.1.3 Είδη Συναρτήσεων Σύζευξης

Οι συναρτήσεις σύζευξης (copulas) χωρίζονται σε τρεις βασικές κατηγορίες, τις οποίες θα παρουσιάσουμε και θα περιγράψουμε στο παρόν κεφάλαιο. Οι τρεις αυτές κατηγορίες αποτελούνται από τις:

- Θεμελιώδεις συναρτήσεις σύζευξης (Fundamental copulas)
- Κλειστής μορφής συναρτήσεις σύζευξης (Explicit copulas)
- Πεπλεγμένης μορφής συναρτήσεις σύζευξης (Implicit copulas).

Θεμελιώδεις συναρτήσεις σύζευξης (Fundamental copulas)

Οι θεμελιώδεις συναρτήσεις σύζευξης (Fundamental copulas) αντιπροσωπεύουν τις τρεις βασικές εξαρτήσεις που μπορεί να έχουν ένα δείγμα από τυχαίες μεταβλητές: 1) η ανεξαρτησία, 2) η τέλεια θετική συσχέτιση και 3) η τέλεια αρνητική συσχέτιση. Για κάθε μια από τις τρεις βασικές εξαρτήσεις υπάρχει και μια συνάρτηση σύζευξης (copula) και ονομάζονται αντίστοιχα:

- Copula ανεξαρτησίας ή γινομένου
- Συν – μονοτονική ή copula ελαχίστου,
- Αντί – μονοτονική ή copula μεγίστου.

Αυτές αποτελούν τις θεμελιώδεις συναρτήσεις σύζευξης (Fundamental copulas) και ανήκουν σε μια ευρύτερη οικογένεια που ονομάζεται Fréchet-Höfdding copulas. Για να παρουσιάσουμε με κατάλληλα παραδείγματα τις τρεις συναρτήσεις σύζευξης (copulas) ως θεωρήσουμε δύο τυχαίες μεταβλητές X, Y , όπως ορίστηκαν στο κεφάλαιο 2.1.2.

Ορισμός 2.10: Ορίζουμε ως συνάρτηση σύζευξης (copula) γινομένου, τη $C[u, v] = uv$ και έχουμε:

$$F_{X,Y}(x, y) = C[F_X(x), F_Y(y)] = F_X(x)F_Y(y)$$

ή:

$$P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$$

Η συνάρτηση σύζευξης (copula) γινομένου αντικατοπτρίζει την ανεξαρτησία μεταξύ δύο τυχαίων μεταβλητών X και Y .

Ορισμός 2.11: Η συνάρτηση σύζευξης (copula) ελαχίστου χρησιμοποιείται όταν δύο τυχαίες μεταβλητές παρουσιάζουν τέλεια θετική αλληλεξάρτηση και για τη διδιάστατη περίπτωση αυτή ισούται με $C[u, v] = \min(u, v)$. Άρα έχουμε:

$$C[F_X(x), F_Y(y)] = \min(F_X(x), F_Y(y))$$

$$P(X \leq x, Y \leq y) = \min(P(X \leq x), P(Y \leq y))$$

Ορισμός 2.12: Η συνάρτηση σύζευξης (copula) μεγίστου αντικατοπτρίζει τη σχέση μεταξύ δύο τυχαίων μεταβλητών, οι οποίες παρουσιάζουν τέλεια αρνητική συσχέτιση. Για τη διδιάστατη περίπτωση αυτή ισούται $C[u, v] = \max(u + v - 1, 0)$ και έχουμε:

$$C[F_X(x), F_Y(y)] = \max(F_X(x) + F_Y(y) - 1, 0)$$

$$P(X \leq x, Y \leq y) = \max(P(X \leq x) + P(Y \leq y) - 1, 0)$$

Κλειστής μορφής συναρτήσεις σύζευξης (Explicit copulas)

Οι συναρτήσεις σύζευξης κλειστής μορφής (Explicit copulas) αποτελούν συναρτήσεις που έχουν κλειστού τύπου μορφή και ο υπολογισμός τους είναι άμεσος και σαφής. Οι 4 βασικότερες κλειστής μορφής συναρτήσεις σύζευξης (explicit copulas) που χρησιμοποιούνται περιλαμβάνουν τις: 1) Gumbel copulas, 2) Clayton copulas, 3) Frank copulas και 4) Αρχιμήδειες copulas. Παρακάτω παρουσιάζονται οι μορφές των συναρτήσεων σύζευξης κλειστής μορφής που αναφέραμε στη διδιάστατη περίπτωση. Σημειώνεται ότι η παράμετρος α στους τύπους έχει γνωστή τιμή και μπορούμε μέσω αυτής να επηρεάσουμε το μέγεθος της συσχέτισης μεταξύ των μεταβλητών.

Ορισμός 2.13: Η Gumbel copula ορίζεται ως εξής:

$$C[u, v] = \exp\left\{-\left((\ln u)^\alpha + (\ln v)^\alpha\right)^{\frac{1}{\alpha}}\right\}$$

Περιγράφει τη δομή εξάρτησης μεταξύ δύο τυχαίων μεταβλητών, στην οποία υπάρχει εξάρτηση άνω ουράς (το επίπεδο της οποίας καθορίζεται από την παράμετρο α) αλλά δεν υπάρχει εξάρτηση κάτω ουράς. Με την αύξηση της παραμέτρου α , αυξάνεται και ο βαθμός εξάρτησης άνω ουράς για την Gumbel copula.

Ορισμός 2.14: Η Clayton copula στη διδιάστατη περίπτωση ορίζεται ως:

$$C[u, v] = (u^{-\alpha} + v^{-\alpha} - 1)^{1/\alpha}$$

Περιγράφει τη δομή της εξάρτησης μεταξύ δύο τυχαίων μεταβλητών, στην οποία υπάρχει εξάρτηση κάτω ουράς (το επίπεδο της οποίας καθορίζεται από την παράμετρο α) αλλά δεν υπάρχει εξάρτηση άνω ουράς.

Ορισμός 2.15: Η Frank copula στη διδιάστατη περίπτωση, περιγράφει τη δομή εξάρτησης μεταξύ δύο τυχαίων μεταβλητών, στην οποία δεν υπάρχει εξάρτηση ούτε άνω, ούτε κάτω ουράς και ορίζεται ως:

$$C[u, v] = -\frac{1}{\alpha} \ln \left(1 + \frac{(e^{-\alpha u} - 1)(e^{-\alpha v} - 1)}{(e^{-\alpha} - 1)} \right)$$

Ορισμός 2.16: Οι Αρχιμήδειες copulas αποτελούν συναρτήσεις σύζευξης (copulas), οι οποίες μπορούν να περιγραφούν με μια γεννήτρια συνάρτηση και η γενική τους μορφή στη διδιάστατη περίπτωση δίνεται από τον παρακάτω τύπο:

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

όπου $\psi(x)$ η γεννήτρια συνάρτηση και $\psi^{[-1]}$ η ψευδό - αντίστροφη συνάρτηση της συνάρτησης ψ . Η γεννήτρια συνάρτηση $\psi: [0,1] \rightarrow (0, \infty]$ πρέπει να είναι συνεχής, γνησίως φθίνουσα και κυρτή με $\psi(1) = 0$.

Ορισμός 2.17: Η ψευδό - αντίστροφη συνάρτηση της $\psi(x)$ ορίζεται ως εξής:

$$\psi^{[-1]}(x) = \begin{cases} \psi^{-1}(x), & 0 \leq x \leq \psi(0) \\ 0 & , \quad \psi(0) \leq x \leq \infty \end{cases}$$

όπου $\psi^{-1}(x)$ δηλώνει τη συνήθη αντίστροφη συνάρτηση που προκύπτει από την αντιστροφή της εξίσωσης $y = \psi(x)$ για να εκφραστεί το x ως προς y .

Παρατήρηση 2.5: Όλες οι προαναφερθείσες συναρτήσεις σύζευξης κλειστής μορφής είναι και Αρχιμήδειες. Για παράδειγμα η Gumbel copula αποτελεί Αρχιμήδεια με γεννήτρια συνάρτηση την $\psi(x) = (-\ln x)^a$ και αντίστροφη την $\psi^{[-1]} = \exp(-x^{\frac{1}{a}})$.

Ορισμός 2.18: Οι Αρχιμήδειες copulas μπορούν να επεκταθούν στην πολυμεταβλητή περίπτωση και έχουν ως γενική μορφή την:

$$C[u_1, \dots, u_d] = \psi^{[-1]} \left(\sum_{i=1}^d \psi(u_i) \right)$$

Επιπλέον, σημειώνεται ότι ο ίδιος ο ορισμός της ψευδό – αντίστροφης συνάρτησης διασφαλίζει ότι η τιμή της Αρχιμήδειας copula, ανεξαρτήτου τιμής του αθροίσματος $\sum_{i=1}^d \psi(u_i)$, θα είναι μια έγκυρη τιμή πιθανότητας.

Πεπλεγμένης μορφής συναρτήσεις σύζευξης (Implicit copulas)

Οι Πεπλεγμένης μορφής συναρτήσεις σύζευξης (Implicit copulas) αποτελούν συναρτήσεις, οι οποίες βασίζονται σε γνωστές πολυμεταβλητές κατανομές, ωστόσο δεν διαθέτουν απλές και αναλυτικές εκφράσεις για την περιγραφή τους. Παρακάτω θα περιγράψουμε την Γκαουσιανή (ή κανονική) copula, η οποία βασίζεται στην πολυμεταβλητή κανονική κατανομή, και τη Student's t copula, η οποία βασίζεται στην πολυμεταβλητή κατανομή του Student t.

Ορισμός 2.18: Η διδιάστατη περίπτωση της Γκαουσιανής copula ορίζεται ως:

$$C[u, v] = \Phi_\rho[\Phi^{-1}(u), \Phi^{-1}(v)]$$

όπου Φ είναι η συνάρτηση κατανομής της τυπικής κανονικής κατανομής και Φ_ρ είναι η συνάρτηση κατανομής της δυσδιάστατης τυπικής κανονικής κατανομής με συσχέτιση μεταξύ των μεταβλητών ρ .

Οι θεμελιώδεις συναρτήσεις σύζευξης (Fundamental copulas) που αναφέραμε προηγουμένως (γινομένου, ελαχίστου και μεγίστου) αποτελούν ειδικές περιπτώσεις της γενικής Γκαουσιανής copula με $\rho = 0$, $\rho = 1$ και $\rho = -1$ αντίστοιχα. Για την πολυμεταβλητή περίπτωση, η Γκαουσιανή copula περιλαμβάνει τον πίνακα συσχέτισης κάθε ζεύγους μεταβλητών ξεχωριστά, μεγέθους $d \times d$. Με αυτόν τον τρόπο παράγει μια από κοινού συνάρτηση κατανομής με έναν συγκεκριμένο πίνακα συσχέτισης από τις μεμονωμένες οριακές κατανομές. Η γενική εξίσωση η οποία υπολογίζει τη διδιάστατη συνάρτηση σύζευξης (copula) ορίζεται ως εξής:

$$C[u, v] = \int_0^u \Phi\left(\frac{\Phi^{-1}(v) - \rho\Phi^{-1}(t)}{\sqrt{1 - \rho^2}}\right) dt$$

Ορισμός 2.19: Η Student's t copula ορίζεται ως:

$$C[u, v] = t_{\gamma, \rho}[t_\gamma^{-1}(u), t_\gamma^{-1}(v)]$$

όπου t_γ είναι μια συνάρτηση κατανομής μιας τυχαίας μεταβλητής με κατανομή Student's t και γ βαθμούς ελευθερίας και $t_{\gamma,\rho}$ είναι η συνάρτηση κατανομής μιας διδιάστατης κατανομής Student's t με συσχέτιση ρ .

Παρατήρηση 2.6: Η Student's t copula επιτρέπει την καλύτερη προσαρμογή των εξαρτήσεων μεταξύ των μεταβλητών και τον υπολογισμό τους με μεγαλύτερη ακρίβεια από ότι η αντίστοιχη Γκαουσιανή, διότι διαθέτει δύο παραμέτρους (γ, ρ αντί μόνο για ρ), με τις οποίες μπορούμε να οριοθετήσουμε τη συσχέτιση και την εξάρτηση μεταξύ των μεταβλητών. Συγκεκριμένα, όσο μειώνεται η παράμετρος γ , τόσο μεγαλύτερη είναι και η εξάρτηση ουράς μεταξύ των μεταβλητών. Επιπρόσθετα, όπως η κανονική κατανομή είναι οριακή περίπτωση της Student's t , έτσι και οι αντίστοιχες copulas έχουν την ίδια συμπεριφορά, καθώς η Student's t τείνει στην Γκαουσιανή, όταν οι τιμές της παραμέτρου γ τείνουν στο άπειρο.

Περαιτέρω πληροφορίες για τα διάφορα μέτρα εξάρτησης καθώς και για τις συναρτήσεις σύζευξης, ο ενδιαφερόμενος αναγνώστης μπορεί π.χ. να ανατρέξει στα συγγράμματα των Cherubini U. & Luciano E. & Vecchiato W. (2004), Joe H. (1997), ActEd (2019), Dupuis D. J. & Jones B. L. (2006).

2.2 ΠΟΛΥΜΕΤΑΒΛΗΤΗ ΘΕΩΡΙΑ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

Στην παρούσα ενότητα εξετάζουμε τα προβλήματα που προκύπτουν όταν επιχειρούμε να μοντελοποιήσουμε την ακραία συμπεριφορά δύο ή περισσότερων (εξαρτημένων) μεταβλητών ταυτόχρονα. Υπάρχουν διάφοροι λόγοι για τους οποίους μπορεί να επιθυμούμε να το κάνουμε αυτό, όπως για τη μοντελοποίηση της κοινής συμπεριφοράς των ακραίων τιμών που εμφανίζονται ως διαδοχικές παρατηρήσεις σε χρονοσειρές (π.χ. ημερήσιες αποδόσεις χρηματιστηριακών τίτλων). Τα προβλήματα αυτά υποδηλώνουν την ανάγκη προσαρμογής μιας κατάλληλης οριακής πολυμεταβλητής κατανομής στα σχετικά δεδομένα. Ωστόσο, όπως θα δούμε, η κατασκευή μιας τέτοιας πολυμεταβλητής κατανομής δεν είναι τόσο εύκολη. Αυξημένη πολυπλοκότητα παρουσιάζεται κατά τη μετάβαση από 1 σε 2 διαστάσεις, οπότε όπως θα αναλυθεί και στο τελευταίο κεφάλαιο, θα επικεντρωθούμε σε μεγάλο βαθμό σε διμεταβλητά προβλήματα και με ποιον τρόπο μπορούμε να τα μοντελοποιήσουμε με ακρίβεια.

Όπως έχουμε περιγράψει και στην αρχή του κεφαλαίου αυτού, θα ασχοληθούμε με τη μετάβαση από τη μονομεταβλητή στην πολυμεταβλητή θεωρία ακραίων τιμών μέσω των συναρτήσεων copulas όπως εκείνες αναλύθηκαν παραπάνω και πιο συγκεκριμένα μέσω της Extreme Value Copula. Και σε αυτή την περίπτωση, υπάρχουν δύο μέθοδοι υπολογισμού των ακραίων τιμών, η μέθοδος Multivariate Block maxima και η μέθοδος Multivariate Peaks over Threshold. Για επιπλέον πληροφορίες επί του αντικειμένου της πολυμεταβλητής θεωρίας ακραίων τιμών, ο ενδιαφερόμενος αναγνώστης μπορεί να ανατρέξει στις εργασίες των Fawcett L. & Walshaw D. (2008), Huang W. (2014), Rocco M. (2011).

Στόχος μας είναι να βρεθεί η από κοινού συνάρτηση κατανομής του διανύσματος των κανονικοποιημένων μεγίστων με τη βοήθεια της Extreme Value Copula και των περιθώριων συναρτήσεων κατανομών που όπως περιγράψαμε στο πρώτο κεφάλαιο συγκλίνουν σε δύο συγκεκριμένες οικογένειες κατανομών: τη γενικευμένη κατανομή ακραίων τιμών (GEV) και τη γενικευμένη κατανομή Pareto (GPD).

2.2.1 Πολυμεταβλητή Μέθοδος Block Maxima

Ας υποθέσουμε ότι θέλουμε να μελετήσουμε την κοινή συμπεριφορά ενός χαρτοφυλακίου j τίτλων.

Τέτοια ζητήματα παρουσιάζουν μεγάλο ενδιαφέρον, ιδίως στις μέρες μας, όπου οικονομικές κρίσεις σε συνδυασμό με ακραία φυσικά φαινόμενα δημιουργούν αβεβαιότητα στην ανθρωπότητα και κατά συνέπεια οδηγούν σε απότομες ακραίες αρνητικές αποδόσεις στα χρηματιστήρια ανά τον κόσμο. Ας υποθέσουμε ότι έχουμε δεδομένα ημερήσιων αποδόσεων για n χρηματιστηριακούς τίτλους. Σε αυτούς ενδέχεται να υπάρχει έντονη εξάρτηση στα άκρα των κατανομών τους, με την έννοια ότι παραδείγματος χάρη μια φυσική καταστροφή μπορεί να επηρεάσει αρνητικά και τους j τίτλους. Οι ακατέργαστες πολυμεταβλητές παρατηρήσεις είναι j -διανύσματα της ημερήσιας απόδοσης κάθε τίτλου.

Η μέθοδος Block Maxima και για την πολυμεταβλητή περίπτωση διαθέτει παρόμοια διαδικασία με τη μονομεταβλητή περίπτωση, δηλαδή χωρίζουμε το δείγμα μας σε m ομάδες και βρίσκουμε το μέγιστο από κάθε ομάδα, το οποίο έχει τη μορφή διανύσματος. Έστω ένα τέτοιο τυχαίο δείγμα με μεταβλητές $\mathbf{X}_1, \dots, \mathbf{X}_n$ με αντίστοιχες περιθώριες συναρτήσεις κατανομής F_1, \dots, F_d . Για κάθε υποομάδα και για κάθε παράγοντα υπάρχει ένα διάνυσμα $M_{n,j} = \max(X_{1,j}, \dots, X_{n,j})$ όπου ο υποδείκτης j δηλώνει τον παράγοντα ($j = 1, \dots, d$).

Τελικός στόχος της ανάλυσης είναι η μοντελοποίηση του τυχαίου διανύσματος $\mathbf{M}_n = (M_{n,1}, \dots, M_{n,d})$. Με την ίδια λογική του Κεντρικού Οριακού Θεωρήματος, μπορούμε να σχηματίσουμε μαθηματικές εκφράσεις (κανονικοποίηση), όπου οι παράμετροι:

$$\mathbf{b}_n = (b_{n,1}, \dots, b_{n,d}) \text{ και } \mathbf{c}_n = (c_{n,1}, \dots, c_{n,d})$$

(το διάνυσμα \mathbf{b}_n πρέπει να είναι αυστηρά θετικό) αποτελούν δύο διανύσματα με γενικό τύπο τον εξής:

$$\frac{\mathbf{M}_n - \mathbf{c}_n}{\mathbf{b}_n} = \left(\frac{M_{n,1} - c_{n,1}}{b_{n,1}}, \dots, \frac{M_{n,d} - c_{n,d}}{b_{n,d}} \right)$$

Η από κοινού συνάρτηση κατανομής για το διάνυσμα $\frac{\mathbf{M}_n - \mathbf{c}_n}{\mathbf{b}_n}$ καθώς το $n \rightarrow \infty$ ονομάζεται Multivariate Extreme Value (MEV) και αποτελεί γενίκευση της μονομεταβλητής Generalized Extreme Value ο τύπος της οποίας παρουσιάζεται παρακάτω ($\mathbf{x} = (x_1, x_2, \dots, x_d)$):

$$\lim_{n \rightarrow \infty} P \left(\frac{\mathbf{M}_n - \mathbf{c}_n}{\mathbf{b}_n} \leq \mathbf{x} \right) = \lim_{n \rightarrow \infty} F^n(\mathbf{b}_n \cdot \mathbf{x} + \mathbf{c}_n) = H(\mathbf{x})$$

Σύμφωνα με το ΘΑΤ, υπάρχουν και μπορούμε να βρούμε τιμές για τις παραμέτρους $\mathbf{b}_n, \mathbf{c}_n$ για τις περισσότερες κατανομές που χρησιμοποιούνται. Γενικεύοντας όσα έχουμε περιγράψει παραπάνω, για οποιαδήποτε κατανομή των αρχικών δεδομένων της ανάλυσης μας, οι κανονικοποιημένες μέγιστες τιμές, όσο το n τείνει στο άπειρο, θα συγκλίνουν σε μία κατανομή που ονομάζεται γενικευμένη κατανομή οριακών τιμών (GEV) με τύπο:

$$\lim_{n \rightarrow \infty} [F(\mathbf{b}_n \mathbf{x} + \mathbf{c}_n)]^n = H(\mathbf{x})$$

Από τη μονοδιάστατη ΘΑΤ προκύπτει ότι οι περιθώριες σ.κ. της πολυδιάστατης $H(\mathbf{x})$ θα πρέπει να είναι μονοδιάστατη GEV.

Επανερχόμενοι στην παραπάνω παρατήρηση, αν υπάρχουν ακολουθίες $\mathbf{b}_n, \mathbf{c}_n$ και οριακή πολυδιάστατη κατανομή $H(\mathbf{x})$ έτσι ώστε:

$$\lim_{n \rightarrow \infty} P \left(\frac{\mathbf{M}_n - \mathbf{c}_n}{\mathbf{b}_n} \leq \mathbf{x} \right) = \lim_{n \rightarrow \infty} F^n(\mathbf{b}_n \cdot \mathbf{x} + \mathbf{c}_n) = H(\mathbf{x})$$

τότε αυτή θα είναι της μορφής:

$$H(\mathbf{x}) = C(G_1(x_1), \dots, G_d(x_d))$$

όπου G_1, \dots, G_d είναι σ.κ. της GEV (με γενικά διαφορετικές παραμέτρους) και C θα είναι μια συνάρτηση σύζευξης η οποία καλείται συνάρτηση σύζευξης ακραίων τιμών (extreme value copula).

Η παραπάνω σχέση για την πολυδιάστατη περίπτωση (u_1, u_2, \dots, u_d) αποτελεί την ικανή και αναγκαία συνθήκη για να θεωρείται μια copula extreme value, και παίρνει την εξής μορφή (Malevergne Y. & Sornette D. 2006):

$$C(u_1, \dots, u_d) = \exp\left(-V\left(-\frac{1}{\ln u_1}, \dots, -\frac{1}{\ln u_d}\right)\right)$$

Η συνάρτηση V έχει τύπο:

$$V(x_1, x_2, \dots, x_d)(w) = \int_{S_d} \max\left(\frac{w_1}{x_1}, \dots, \frac{w_d}{x_d}\right) dH(w_1, \dots, w_d)$$

όπου το H είναι ένα πεπερασμένο μέτρο τέτοιο ώστε $\int_{S_d} w_i dH(w_1, \dots, w_d) = 1$ και το S_d ισούται με:

$$S_d = \left\{ (w_1, w_2, \dots, w_d) \in \mathbb{R}_+^d : \sum_{i=1}^d w_i = 1, \forall x_i > 0 \right\}$$

Για την εκτίμηση των παραμέτρων της κάθε extreme value copula μπορούν να χρησιμοποιηθούν πολλοί τρόποι, με έναν από αυτούς να είναι ο παραμετρικός τρόπος εκτίμησης της μέγιστης πιθανοφάνειας. Ο Joe H. (1997) ανέλυσε τον παραμετρικό τρόπο εκτίμησης της μέγιστης πιθανοφάνειας και απέδειξε ότι αυτός πρακτικά επικεντρώνεται στη μεγιστοποίηση της ακόλουθης συνάρτησης:

$$l(\theta) = \sum_{t=1}^n \ln C(F_1(x_{1,t}; \theta_1), F_2(x_{2,t}; \theta_2); \theta) + \sum_{t=1}^n \left(\ln(f_1(x_{1,t}; \theta_1), f_2(x_{2,t}; \theta_2)) \right)$$

Εάν παραγωγίζεται, η μεγιστοποίησή της πραγματοποιείται εξισώνοντας την παράγωγό της με το μηδέν. Γενικά η εκτιμήτρια μέγιστης πιθανοφάνειας των παραμέτρων είναι:

$$\hat{\theta} = \operatorname{argmax} l(\theta)$$

Η πολυπλοκότητα της συνάρτησης αυτής, οδήγησε τους Shih J.H. & Louis T.A. (1995) και Joe H. & Xu J. (1996) να αναπτύξουν μια εναλλακτική συνάρτηση την inference function for margin (IMF), κατά την οποία η μεγιστοποίηση της παραπάνω συνάρτησης πραγματοποιείται δοθέντος των περιθώριων συναρτήσεων μέσα σε δύο φάσεις. Αρχικά εκτιμώνται οι παράμετροι της περιθώριας κατανομής και στη συνέχεια οι αντίστοιχες της copula. Το μεγαλύτερο πλεονέκτημα της προσέγγισης αυτής έγκειται στο γεγονός ότι μπορεί να εφαρμοστεί ευκολότερα σε διανύσματα πολλαπλών διαστάσεων αλλά και στις ιδιότητες της συνέπειας και της ασυμπτωματικής κανονικότητας που έχει.

2.2.2 Πολυμεταβλητή Μέθοδος Peaks over Threshold

Όπως και στη μονομεταβλητή περίπτωση έτσι και στην πολυμεταβλητή η κατανομή η οποία ουσιαστικά παίζει ρόλο στην ανάλυση της μεθόδου είναι η Generalized Pareto και το κατώφλι u θα είναι ένα διάνυσμα της μορφής $\mathbf{u} = (u_1, \dots, u_d)$. Οι παραμετρικές μέθοδοι μοντελοποιούν ένα τυχαίο πολυδιάστατο δείγμα με συνάρτηση κατανομής F που πληροί την ιδιότητα του Domain of attraction για τη MEV. Επομένως, έστω ένα διάνυσμα $M_n = (M_n^{(1)}, \dots, M_n^{(d)}) = (\max(X_i^{(1)}), \dots, \max(X_i^{(d)})) \forall i = 1, \dots, d$. Χρησιμοποιώντας την ίδια λογική όπως και στη μέθοδο block maxima, για το διάνυσμα των μέγιστων τιμών για κάθε διάνυσμα θα ισχύει ότι:

$$P \left[\frac{M_n^{(i)} - \beta_n^{(i)}}{\alpha_n^{(i)}} \leq x^i, 1 \leq i \leq d \right] = F^n(a_n^{(1)}x^1 + \beta_n^{(1)}, \dots, a_n^{(d)}x^d + \beta_n^{(d)}) \rightarrow G(x)$$

Η $G(x)$ αποτελεί μια extreme value κατανομή, με μορφή που φαίνεται παρακάτω:

$$G(x) = \begin{cases} \exp\left(-\left(1 + \gamma_i \frac{x - \mu_i}{\sigma_i}\right)^{-\frac{1}{\gamma_i}}\right), & \gamma_i \neq 0 \\ \exp\left(-\exp\left(-\frac{x - \mu_i}{\sigma_i}\right)\right), & \gamma_i = 0 \end{cases}$$

Για την περίπτωση της μεθόδου Peaks over Threshold η χρήση copulas είναι εξίσου ευρέως διαδεδομένη και η αντίστοιχη copula για τη γενικευμένη κατανομή Pareto να έχει την εξής σχέση:

$$F(x) = C_0(F_1(x_1), \dots, F_d(x_d)), \quad x \geq u$$

με τη C_0 να αποτελεί συνάρτηση σύζευξης ακραίων τιμών (Extreme value copula).

2.2.3 Μοντελοποίηση της Δομής Εξάρτησης

Η κλάση των διμεταβλητών μοντέλων ακραίων τιμών περιλαμβάνει πολλές οικογένειες κατανομών που μπορούν να χρησιμοποιηθούν για τη μοντελοποίηση της δομής εξάρτησης στα δεδομένα.

Παραπάνω δόθηκε ήδη η γενική μορφή που πρέπει να διαθέτει μια συνάρτηση σύζευξης ακραίων τιμών. Μια ισοδύναμη αναπαράσταση της οριακής κατανομής των διδιάστατων block maxima αποδεικνύεται ότι είναι:

$$C(z_1, z_2) = \exp \left\{ -(y_1 + y_2) A \left(\frac{y_1}{y_1 + y_2} \right) \right\}$$

όπου:

$$y_j = y_j(z_j) = \left\{ 1 + \frac{\xi_j(z_j - \mu_j)}{\sigma_j} \right\}_+^{-\frac{1}{\xi_j}}$$

για $\sigma_j > 0$ και $j = 1, 2, \dots$. Η συνάρτηση:

$$A(x) = -\log\{G(y_1^{-1}(x), y_2^{-1}(1-x))\}$$

που ορίζεται για $0 \leq x \leq 1$ καλείται συνάρτηση εξάρτησης. Οι οριακές περιθώριες συναρτήσεις είναι GEV, και είναι της μορφής $G_j(z_j) = \exp(-y_j)$. Συνεπάγεται ότι $A(0) = A(1) = 1$ και ότι η A είναι κυρτή συνάρτηση με $\max(x, 1-x) \leq A(x) \leq 1$ για όλα τα $0 \leq x \leq 1$. Σε πλήρη ανεξαρτησία $A(1/2) = 1$ ενώ σε πλήρη εξάρτηση $A(1/2) = 0.5$

Υπάρχουν διάφορα παραμετρικά διμεταβλητά μοντέλα ακραία τιμών για τον υπολογισμό της από κοινού συνάρτησης κατανομής των κανονικοποιημένων μεγίστων από ακραίες παρατηρήσεις. Μερικά από αυτά παρουσιάζονται παρακάτω.

Το λογιστικό (log) μοντέλο (Gumbel, 1960)

Η διμεταβλητή λογιστική συνάρτηση κατανομής με παράμετρο $dep = r$ είναι:

$$G(z_1, z_2) = C(G_1(z_1), G_2(z_2)) = \exp[-(y_1^{1/r} + y_2^{1/r})^r]$$

όπου:

$$C(x_1, x_2) = \exp(-((- \ln x_1)^{1/r} + (- \ln x_2)^{1/r})^r)$$

με $0 < r \leq 1$. Πρόκειται για μια ειδική περίπτωση του διμεταβλητού ασύμμετρου λογιστικού μοντέλου. Πλήρης εξάρτηση επιτυγχάνεται στο όριο όταν το r πλησιάζει το μηδέν. Ανεξαρτησία επιτυγχάνεται όταν $r = 1$.

Το ασύμμετρο λογιστικό (alog) μοντέλο (Tawn, 1988)

Η διμεταβλητή ασύμμετρη λογιστική συνάρτηση κατανομής με παραμέτρους $dep = r$ και $asy = (t_1, t_2)$ είναι:

$$G(z_1, z_2) = \exp\{-(1-t_1)y_1 - (1-t_2)y_2 - [(t_1y_1)^{1/r} + (t_2y_2)^{1/r}]^r\}$$

όπου $0 < r \leq 1$ και $0 \leq t_1, t_2 \leq 1$. Όταν $t_1 = t_2 = 1$ το ασύμμετρο λογιστικό μοντέλο είναι ισοδύναμο με το λογιστικό μοντέλο. Ανεξαρτησία επιτυγχάνεται όταν είτε $r = 1, t_1 = 0$ ή $t_2 = 0$. Πλήρης εξάρτηση επιτυγχάνεται στο όριο όταν $t_1 = t_2 = 1$ και το r πλησιάζει το μηδέν. Διαφορετικά όρια εμφανίζονται όταν t_1 και t_2 είναι σταθερά και το r πλησιάζει το μηδέν.

Το μοντέλο Husler-Reiss (Husler και Reiss, 1989)

Η συνάρτηση κατανομής Husler-Reiss με παράμετρο $dep = r$ είναι:

$$G(z_1, z_2) = \exp\left(-y_1 \Phi\left\{r^{-1} + \frac{1}{2}r[\log(y_1/y_2)]\right\} - y_2 \Phi\left\{r^{-1} + \frac{1}{2}r[\log(y_2/y_1)]\right\}\right)$$

όπου $\Phi(\cdot)$ είναι η τυπική κανονική συνάρτηση κατανομής και $r > 0$. Η ανεξαρτησία επιτυγχάνεται στο όριο καθώς το r πλησιάζει το μηδέν. Πλήρης εξάρτηση επιτυγχάνεται καθώς το r τείνει στο άπειρο.

Το αρνητικό λογιστικό (neglog) μοντέλο (Galambos, 1975)

Η διμεταβλητή αρνητική λογιστική συνάρτηση κατανομής με παράμετρο $dep = r$ είναι:

$$G(z_1, z_2) = \exp\{-y_1 - y_2 + [y_1^{-r} + y_2^{-r}]^{-1/r}\}$$

όπου $r > 0$. Πρόκειται για μια ειδική περίπτωση του διμεταβλητού ασύμμετρου αρνητικού λογιστικού μοντέλου. Η ανεξαρτησία επιτυγχάνεται στο όριο καθώς το r πλησιάζει το μηδέν. Πλήρης εξάρτηση επιτυγχάνεται καθώς το r τείνει στο άπειρο.

Το ασύμμετρο αρνητικό λογιστικό (anelog) μοντέλο (Joe, 1990)

Η διμεταβλητή ασύμμετρη αρνητική λογιστική συνάρτηση κατανομής με παραμέτρους $dep = r$ και $asy = (t_1, t_2)$ είναι:

$$G(z_1, z_2) = \exp\{-y_1 - y_2 + [(t_1 y_1)^{-r} + (t_2 y_2)^{-r}]^{-1/r}\}$$

όπου $r > 0$ και $0 < t_1, t_2 \leq 1$. Όταν $t_1 = t_2 = 1$ το ασύμμετρο αρνητικό λογιστικό μοντέλο είναι ισοδύναμο με το αρνητικό λογιστικό μοντέλο. Η ανεξαρτησία επιτυγχάνεται στο όριο όταν είτε r, t_1 ή t_2 πλησιάζει το μηδέν. Πλήρης εξάρτηση επιτυγχάνεται στο όριο όταν $t_1 = t_2 = 1$ και το r τείνει στο άπειρο. Διαφορετικά όρια εμφανίζονται όταν τα t_1 και t_2 είναι σταθερά και το r τείνει στο άπειρο.

Για μια πληρέστερη περιγραφή των διμεταβλητών μοντέλων ακραίων τιμών, ο ενδιαφερόμενος αναγνώστης μπορεί να ανατρέξει στις εργασίες των Bouyé E. (2002), Stephenson A. G. (2003), Stephenson, A. G. (2018).

ΚΕΦΑΛΑΙΟ 3ο:

ΕΠΙΜΕΤΡΗΣΗ ΚΙΝΔΥΝΟΥ

3.1 ΜΕΤΡΑ ΚΙΝΔΥΝΟΥ

Η τεχνική μέτρησης του κινδύνου μέσω της μεθόδου VaR αποτελεί ένα στατιστικό εργαλείο το οποίο περιγράφει τις συνολικές πιθανές απώλειες ενός χαρτοφυλακίου. Πιο συγκεκριμένα, μετράει τις απώλειες λόγω των λεγόμενων «κανονικών» μετατοπίσεων των επενδυτικών αγορών. Οι απώλειες οι οποίες ξεπερνούν την τιμή VaR έχουν μια μικρή πιθανότητα να συμβούν. Πρακτικά μέσω της τεχνικής αυτής, αθροίζονται όλα τα πιθανά ρίσκα από τις επενδύσεις στο σύνολο ενός χαρτοφυλακίου και προκύπτει ένας μοναδικός αριθμός, ο οποίος και περιγράφει τις συνολικές πιθανές απώλειες.

Όσον αφορά τον τεχνικό ορισμό της VaR μπορούμε να την περιγράψουμε ως εξής: χρησιμοποιώντας μια πιθανότητα τιμής x επί τις εκατό και μια διάρκεια κατοχής t σε ημέρες, η VaR ορίζεται ως η απώλεια η οποία αναμένεται να ξεπεραστεί με πιθανότητα υπέρβασης x τις επόμενες t ημέρες κατοχής. Τυπικές τιμές πιθανοτήτων υπέρβασης x είναι 1, 2.5 και 5 τις εκατό, ενώ τυπικές τιμές για τη διάρκεια κατοχής t είναι 1, 2 και 10 (εργάσιμες) μέρες καθώς και ένας μήνας. Η επιλογή της τιμής x εναπόκειται στον επενδυτή και εξαρτάται από τον τρόπο χαρακτηρισμού των απωλειών ως σημαντικές ή μη. Παράλληλα, οι χρόνοι κατοχής προκύπτουν σύμφωνα με τον επενδυτικό ορίζοντα, με τους πιο δραστήριους επενδυτές να χρησιμοποιούν διάρκεια κατοχής τη μία ημέρα, ενώ άλλοι μεγαλύτερες περιόδους.

Εδώ θα πρέπει να σημειωθεί ότι η τεχνική προϋποθέτει ότι το χαρτοφυλάκιο θα παραμείνει σταθερό και χωρίς αλλαγές κατά τη διάρκεια κατοχής, κάτι το οποίο δεν είναι ρεαλιστικό ειδικότερα για μεγαλύτερη διάρκεια t . Επιπλέον, έχουν παρατηρηθεί ακόμη μερικά προβλήματα στη χρήση της τεχνικής VaR με τα βασικότερα από αυτά να είναι: 1) η VaR μετράει μόνο τα ποσοστημόρια των κατανομών κέρδους – ζημίας και δεν παρέχει καμία πληροφορία μετά την οριακή τιμή VaR –το φαινόμενο αυτό ονομάζεται «κίνδυνος ουράς», και 2) η VaR δεν αποτελεί συνεκτικό μέτρο κινδύνου (coherent risk measure) διότι τα επιμέρους ρίσκα από όλες τις επενδύσεις του χαρτοφυλακίου δεν ισούνται με

την τιμή VaR. Επίσης, σε χαρτοφυλάκια που χαρακτηρίζονται από κατανομές με μεγάλες ουρές και αυξημένες πιθανότητες για υψηλές απώλειες, η τεχνική VaR υποεκτιμά το ρίσκο. Για αυτόν τον λόγο μπορεί να παρουσιάζονται επενδύσεις με σημαντικότερες πιθανότητες μεγάλων ζημιών ως λιγότερο ριψοκίνδυνες από άλλες.

Τα παραπάνω μειονεκτήματα της τεχνικής VaR, οδήγησαν στην ανάπτυξη και χρήση μιας άλλης τεχνικής, του αναμενόμενου ελλείματος (expected shortfall, ES). Η ES ορίζεται ως η αναμενόμενη απώλεια ενός χαρτοφυλακίου στις συνθήκες κατά τις οποίες οι απώλειες έχουν ξεπεράσει την τιμή VaR. Επιπροσθέτως, η ES έχει την ιδιότητα της υποπροσθετικότητας, δηλαδή η τιμή της είναι μικρότερη ή ίση με το άθροισμα των απωλειών των επιμέρους επενδύσεων του χαρτοφυλακίου και χαρακτηρίζεται ως συνεκτικό μέτρο κινδύνου.

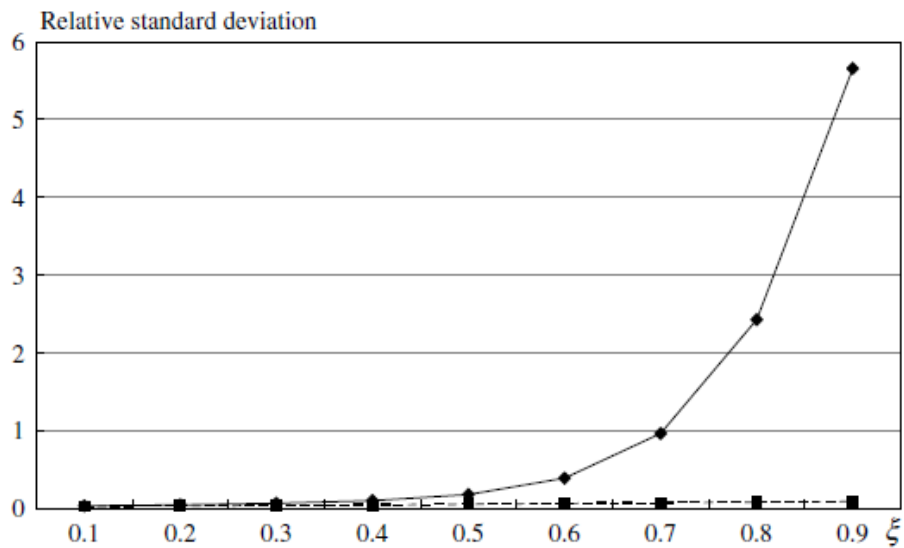
Η ES ουσιαστικά παρουσιάζει τη μέση απώλεια του χαρτοφυλακίου, όταν οι απώλειες ξεπερνούν την τιμή της VaR και ορίζεται σύμφωνα με τον παρακάτω τύπο:

$$ES_{\alpha}(X) = E[X|X \geq VaR_{\alpha}(X)]$$

Σε περιπτώσεις όπου η σχέση κέρδους – ζημίας διαθέτει κανονική συνάρτηση κατανομής, η VaR και η ES έχουν πρακτικώς το ίδιο αποτέλεσμα. Οι σημαντικές διαφορές παρατηρούνται όταν η παραπάνω κατανομή δεν είναι κανονική και έχει βαριά ουρά. Επιπλέον, κατά τη διάρκεια μεγάλων οικονομικών γεγονότων όταν οι αγορές βρίσκονται υπό πίεση και οι επιμέρους επενδύσεις ενός χαρτοφυλακίου έχουν σημαντική διακύμανση η VaR αποτυγχάνει στην ανάλυση και απομένει μόνο η ES.

Όπως αναφέρθηκε, η ES επιδεικνύει καλύτερη συμπεριφορά από τη VaR στις περιπτώσεις μεγάλων ουρών. Ωστόσο αυτό δεν σημαίνει ότι αποφέρει και πιο ορθά αποτελέσματα από τη VaR, ειδικά αν χρησιμοποιηθούν μέθοδοι προσομοίωσης για τις εκτιμήσεις αυτών. Τα αποτελέσματα ενδέχεται να μην είναι ακριβή, αν βασίζονται σε περιορισμένο μέγεθος δειγμάτων για την εκτίμησή τους. Το σφάλμα αυτό μειώνεται με την αύξηση του αριθμού των δειγμάτων τυχαίων αριθμών που παράγονται με τη μέθοδο Monte – Carlo. Το παραπάνω, όπως είναι φυσικό, συνεισφέρει στην άμεση αύξηση του υπολογιστικού χρόνου της ανάλυσης μας. Με κατάλληλη ανάλυση οι Yamai Y. & Yoshihara T. (2005) απέδειξαν ότι για κατανομές απωλειών με μεγάλες ουρές, το σφάλμα της εκτίμησης της ES είναι μεγαλύτερο από αυτό της VaR.

Στο παρακάτω διάγραμμα παρουσιάζεται η σχετική τυπική απόκλιση (τυπική απόκλιση δια τον μέσο όρο) για της ES (συνεχόμενη γραμμή) και της VaR (διακεκομμένη γραμμή). Όσο αυξάνεται η τιμή της παραμέτρου ξ (κατανομές των επιμέρους επενδύσεων με όλο και μεγαλύτερες ουρές) φαίνεται και η διαφορά στο σφάλμα εκτίμησης των δύο μεθόδων.



* Πηγή γραφήματος: Yamai Y. & Yoshida T., (2005), Value-at-risk versus expected shortfall: A practical perspective, Journal of Banking & Finance 29, 997-1015

Μια εξήγηση για την παραπάνω συμπεριφορά είναι ότι, η εκτίμηση της ES εξαρτάται και επηρεάζεται άμεσα από την πιθανότητα ακραίων και μεγάλων απωλειών, εφόσον η ES λαμβάνει υπ'όψιν της τη δεξιά ουρά των κατανομών όπου και εντοπίζονται αυτές οι απώλειες. Αντίθετα, όπως έχει ήδη αναφερθεί οι εκτιμήσεις της VaR δεν λαμβάνουν υπ'όψιν τις απώλειες πάνω από την τιμή της και δεν επηρεάζεται ουσιαστικά από αυτά τα ακραία φαινόμενα απωλειών. Επομένως, όταν η κατανομή απωλειών έχει μεγαλύτερες ουρές ($\xi > 0$), οι εκτιμήσεις της ES έχουν μεγαλύτερη διακύμανση και το σφάλμα μεγαλώνει αντίστοιχα.

Τέλος, με κατάλληλη διερεύνηση με στόχο την ελαχιστοποίηση του σφάλματος και τη σύγκλιση της ES, οι Yamai Y. & Yoshida T. (2005) αυξάνοντας το μέγεθος του δείγματος για τη μέθοδο Monte Carlo από 10 000 σε 1 000 000 κατέληξαν σε αποτελέσματα τα οποία διαθέτουν αποδεκτά σφάλματα.

3.2 ΜΕΘΟΔΟΙ ΥΠΟΛΟΓΙΣΜΟΥ ΤΟΥ VaR

Οι μεθοδολογίες για την εύρεση της VaR περιλαμβάνουν τη μέθοδο Monte Carlo, η οποία αναφέρθηκε και προηγουμένως, τη μέθοδο Variance Covariance, τις προσομοιώσεις με βάση ιστορικά δεδομένα, τα λεγόμενα μοντέλα GARCH, καθώς και στοχαστικά μοντέλα διακύμανσης (SV).

Η προσομοίωση μέσω ιστορικών δεδομένων (εμπειρική εκτίμηση του VaR, ES) αποτελεί μια μέθοδο η οποία δεν διαθέτει πολλές συμβάσεις και παραδοχές για τις στατιστικές κατανομές των παραγόντων που ορίζουν τις αγορές. Πρακτικά, χρησιμοποιούνται οι ιστορικές αλλαγές στα επιτόκια και στις τιμές των αγορών για τη δημιουργία μιας κατανομής των πιθανών μελλοντικών κερδών και απωλειών του χαρτοφυλακίου. Από την κατανομή αυτή, λαμβάνεται και η τιμή VaR ως το όριο το οποίο ξεπερνούν οι απώλειες στο $p\%$ των τιμών.

Μια άλλη μεθοδολογία εύρεσης της VaR είναι η προσομοίωση Monte Carlo, η οποία εμφανίζει αρκετές ομοιότητες με την προσομοίωση με ιστορικά δεδομένα. Η βασική τους διαφορά έγκειται, στο ότι αντί για τη χρήση παρατηρήσεων για τους παράγοντες των αγορών τις τελευταίες N ημέρες, κατά την προσομοίωση Monte Carlo δημιουργούνται τυχαίες τιμές για αυτούς. Αυτό επιτυγχάνεται με την επιλογή μιας κατάλληλης στατιστικής κατανομής η οποία θεωρείται ότι περιγράφει επαρκώς τις πιθανές μελλοντικές αλλαγές στους παράγοντες των αγορών. Στη συνέχεια, παράγονται χιλιάδες τυχαίοι αριθμοί, σύμφωνα και με αυτήν την κατανομή, μέσω των οποίων δημιουργούνται υποθετικά κέρδη και ζημιές για το χαρτοφυλάκιο και η εξάγεται η κατανομή αυτών. Τελικά, η τιμή VaR υπολογίζεται από την παραγόμενη κατανομή κέρδους – ζημίας όπως αναφέρθηκε και στην παραπάνω μεθοδολογία.

Η μέθοδος αυτή χαρακτηρίζεται ως η πιο άμεση καθώς υπολογίζει τη μέγιστη ζημία ενός χρηματοπιστωτικού ιδρύματος σε ένα δεδομένο χρονικό διάστημα και με συγκεκριμένη πιθανότητα. Το χαρακτηριστικό της είναι ότι παρατηρεί τις κινήσεις των στοιχείων που απαρτίζουν το χαρτοφυλάκιο και χρησιμοποιώντας την τυπική απόκλιση των κινήσεων αυτών, υποθέτοντας παράλληλα κανονική κατανομή για τις εκτιμήτριες συναρτήσεις των ποσοτήτων που μας ενδιαφέρουν, υπολογίζει τη μέγιστη απώλεια μέσα σε ένα προκαθορισμένο διάστημα εμπιστοσύνης (συνήθως 95%) βασιζόμενο, όπως γίνεται αντιληπτό, στη μεταβλητότητα των τιμών ή πτητικότητα (volatility) του χαρτοφυλακίου. Ωστόσο, λαμβάνοντας υπόψη τη μεταβλητότητα κάθε στοιχείου ξεχωριστά, δεν είναι

δυνατό να εξαχθούν ασφαλή συμπεράσματα για ολόκληρο το χαρτοφυλάκιο. Κατά αυτόν τον τρόπο, ένα σημείο που θα πρέπει να εξετάζεται είναι και η συσχέτιση που έχουν τα κέρδη ή οι ζημιές των στοιχείων που συγκροτούν ένα χαρτοφυλάκιο. Για παράδειγμα, αν όλα τα στοιχεία αντιδρούν με τον ίδιο τρόπο τότε καταλήγουμε ότι έχουν πλήρη θετική συσχέτιση. Αν τα πράγματα πηγαίνουν καλά, τότε έχουμε ένα θετικό αποτέλεσμα στο χαρτοφυλάκιο μας. Αν όμως μειωθεί η αξία των στοιχείων, τότε τα άσχημα νέα είναι αναπόφευκτα. Αντιθέτως, αν τα στοιχεία δεν είναι θετικά συσχετισμένα, τότε μία θετική κίνηση στο ένα στοιχείο θα αποτελέσει μία άσχημη κίνηση σε ένα άλλο. Συνεπώς, είναι σημαντική η επίγνωση και της συσχέτισης των στοιχείων, κάτι που η μέθοδος Variance Covariance παρέχει.

Το μοντέλο EWMA μοντελοποιεί τις διασπορές και τις συνδιασπορές με τη χρήση ενός εκθετικά κινούμενου μέσου όρου. Στις τιμές των παρατηρήσεων δίνονται διαφορετικές βαρύτητες, με τα πρόσφατα δεδομένα να λαμβάνονται υπόψιν περισσότερο, ενώ όσο κινούμαστε στο παρελθόν τα βάρη μειώνονται ταχύτερα. Επομένως, δίνεται μεγαλύτερη βαρύτητα στο πρόσφατο παρελθόν, ενώ έχει άμεση αντίδραση σε πιθανή κατάρρευση των αγορών. Η εξίσωση της μεταβλητότητας ορίζεται ως εξής:

$$\sigma_t^2 = (1 - \lambda) \sum_{t=1}^n \lambda^2 (R_t - \mu)^2$$

όπου: σ_t^2 είναι η τυπική απόκλιση,

R_t είναι η επενδυτική απόδοση τη στιγμή t ,

μ ο μέσος όρος της απόδοσης,

t ο χρονικό ορίζοντα,

λ ο εκθετικός παράγοντας που δείχνει την επιμονή της μεταβλητότητας της αγοράς (μεταξύ 0 και 1),

$(1 - \lambda)$ η παράμετρος που δείχνει την ταχύτητα με την οποία ένα σοκ της αγοράς απορροφάται από τη μεταβλητότητα.

Η προσέγγιση αυτή έχει δύο βασικά πλεονεκτήματα με πρώτο τη γρηγορότερη απόκριση σε σοκ των αγορών, μιας και τα πιο πρόσφατα δεδομένα έχουν και μεγαλύτερη βαρύτητα από τα παλαιότερα. Δεύτερον, η μεταβλητότητα φθίνει εκθετικά μετά από κάθε σημαντικό σοκ, εφόσον η βαρύτητα του παρατηρούμενου σοκ μειώνεται αντίστοιχα ως λ^{t-1} . Η βασικότερη επιλογή για το μοντέλο είναι αυτή του εκθετικού παράγοντα λ . Εάν

ο αριθμό λ είναι μεγάλος, οι τρέχουσες διακυμάνσεις θα έχουν μικρή επίδραση στη συνολική διακύμανση του μοντέλου. Η τιμή του εξαρτάται από τον χρονικό ορίζοντα του επενδυτή με χαρακτηριστικές τιμές από το RiskMetrics να είναι 0.94 για ημερήσιο και 0.97 για μηνιαίο επενδυτικό χρονικό ορίζοντα.

Μια άλλη προσέγγιση είναι το μοντέλο GARCH, το οποίο χρησιμοποιείται για την εκτίμηση μεταβλητότητας και της συσχέτισης μεταξύ διαφορετικών επενδύσεων. Το απλούστερο GARCH(1,1) είναι το:

$$\sigma_t^2 = \gamma + \beta \cdot \sigma_{t-1}^2 + \alpha \cdot X_{t-1}^2$$

όπου γ , α και β είναι παράμετροι που χρήζουν εκτίμησης. Οι παράμετροι α και β περιγράφουν την επιμονή της μεταβλητότητας και το άθροισμα τους θα πρέπει να ισούται με 1. Υψηλές τιμές α και β υποδεικνύουν και υψηλή μεταβλητότητα, ενώ αντιστοιχούν πρακτικά με τις παραμέτρους λ και $1 - \lambda$ του μοντέλου EWMA. Επιπροσθέτως, εάν η παράμετρος γ ισούται με μηδέν τότε η εξίσωση EWMA είναι μια ειδική εκδοχή του μοντέλου GARCH. Η εξίσωση του μοντέλου GARCH(p, q) είναι:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \cdot X_{t-i}^2 + \sum_{j=1}^q \beta_j \cdot \sigma_{t-j}^2$$

και η απόδοση της επένδυσης καθορίζεται από την επόμενη εξίσωση:

$$R_t = \sigma_t \cdot X_t + \varepsilon_t$$

όπου ε_t είναι λευκός θόρυβος. Οι παράμετροι α_0 , α_i και β_j πρέπει να είναι μεγαλύτερες του μηδενός και επίσης να ικανοποιούν την παρακάτω συνθήκη:

$$0 \leq \sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j \leq 1$$

Το μοντέλο αυτό έχει δύο κύρια στοιχεία: τον ορισμό της εξίσωσης της διακύμανσης και την παραδοχή ότι τα κανονικοποιημένα σφάλματα έχουν τη μορφή του «τυχαίου περιπάτου».

Για να εκτιμήσουμε την τιμή VaR χρησιμοποιώντας το μοντέλο GARCH, δημιουργούμε την κατανομή των ημερήσιων τιμών για το χαρτοφυλάκιο και στη συνέχεια να προσαρμόσουμε το μοντέλο σε αυτές με τις κατάλληλες τιμές των παραμέτρων α, β, γ . Έπειτα, πραγματοποιείται ένας μεγάλος αριθμός προσομοιώσεων τόσων ημερών, όσο και ο μέγιστος επενδυτικός χρονικός ορίζοντας. Μετά το τέλος των προσομοιώσεων μπορούμε

να βρούμε τα επιλεγμένα εκατοστημόρια της κατανομής του χαρτοφυλακίου για κάθε προσομοιωμένη ημέρα.

Όπως όλα τα μοντέλα έτσι και αυτό αποτελείται από διάφορες πτυχές, οι οποίες μπορούν να βελτιωθούν και με αυτόν τον τρόπο να επιτευχθεί ορθότερη περιγραφή των χαρακτηριστικών και της δυναμικής των χρονοσειρών που εξετάζονται. Στη βάση αυτή, υπάρχουν διάφορες προεκτάσεις του βασικού μοντέλου όπως τα: Asymmetric GARCH, Exponential GARCH, Integrated GARCH και GARCH in Mean μοντέλα, τα οποία έχουν διαφορετική απόκριση στα πλήγματα των αγορών.

3.3 VaR ΣΕ ΕΠΙΠΕΔΟ ΧΑΡΤΟΦΥΛΑΚΙΟΥ

Σε ένα ρεαλιστικό οικονομικό περιβάλλον, η χάραξη στρατηγικής της διαχείρισης του επενδυτικού ρίσκου, πραγματοποιείται σε επίπεδο χαρτοφυλακίου, καθώς η λογική των επενδύσεων και η συνήθης πρακτικής τους, επιτάσσει την ταυτόχρονη επένδυση σε διάφορους τίτλους. Αυτή η πρακτική έχει ως στόχο την εξομάλυνση του συνολικού κινδύνου και απωλειών με την αντιστάθμιση των όποιων αρνητικών επιδόσεων ενός τίτλου από αντίστοιχες θετικές ενός άλλου. Συνεπώς, η γνώση των μέγιστων απωλειών ενός χαρτοφυλακίου, με βάση και μια πιθανότητα p , κρίνεται αναγκαία.

Αν $\mathbf{X} = (X_1, X_2, \dots, X_n)$ το διάνυσμα όλων των επενδυσόμενων θέσεων ενός χαρτοφυλακίου, και οι αντίστοιχες διασπορές είναι οι $\mathbf{Var} = (Var_1, Var_2, \dots, Var_n)$, η διασπορά του συνόλου του χαρτοφυλακίου υπολογίζεται με βάση την παρακάτω γενικευμένη σχέση:

$$\mathbf{Var} = \sqrt{\left(\sum_{i=1}^n Var_i^2 + 2 \sum_{i<j}^n \rho_{ij} Var_i Var_j \right)}$$

με ρ_{ij} : ο δείκτης συσχέτισης μεταξύ των αποδόσεων των i και j επενδυσόμενων τίτλων.

Στην ειδική περίπτωση, κατά την οποία το χαρτοφυλάκιο περιλαμβάνει δύο ή τρεις επενδυσόμενους τίτλους, η παραπάνω σχέση παίρνει τις ακόλουθες απλούστερες μορφές αντίστοιχα:

$$Var = \sqrt{(Var_1^2 + Var_2^2 + 2\rho_{12}Var_1Var_2)}$$

$$Var = \sqrt{(Var_1^2 + Var_2^2 + Var_3^2 + 2\rho_{12}Var_1Var_2 + 2\rho_{13}Var_1Var_3 + 2\rho_{23}Var_2Var_3)}$$

Από τα παραπάνω προκύπτει η ανάγκη για την εκτίμηση και της συσχέτισης μεταξύ των μεταβλητών των διαφορετικών επενδύσεων, πέρα από τη διακύμανσή τους. Ένας τρόπος για την εκτίμηση της συσχέτισης αυτής είναι η χρήση των συναρτήσεων copulas, όπως έχει περιγράψει στο κεφάλαιο 2.

ΚΕΦΑΛΑΙΟ 4ο:

ΠΡΑΚΤΙΚΗ ΕΦΑΡΜΟΓΗ

ΥΠΟΛΟΓΙΣΜΟΥ ΤΟΥ VaR

4.1 ΕΙΣΑΓΩΓΗ

Στο παρόν κεφάλαιο παρουσιάζεται η εφαρμογή των συναρτήσεων σύζευξης σε συνδυασμό με την Πολυμεταβλητή Θεωρία Ακραίων Τιμών, χρησιμοποιώντας πραγματικά δεδομένα δισδιάστατων αποδόσεων. Η ανάλυση θα βασιστεί στη Συνάρτηση Σύζευξης Ακραίων Παρατηρήσεων (Extreme value Copula). Ως εφαρμογή, θα πραγματοποιηθεί μέσω του στατιστικού πακέτου R , η εκτίμηση μέτρων κινδύνου όπως το VaR. Πιο συγκεκριμένα, οι μέθοδοι εφαρμόζονται σε ένα σύνολο πραγματικών δισδιάστατων αποδόσεων που κατασκευάζεται από τα κρυπτονομίσματα του Bitcoin και του Cardano, με σκοπό την εκτίμηση του VaR μέσω της διμεταβλητής (από κοινού) συνάρτησης κατανομής ακραίων παρατηρήσεων (Bivariate extreme value distribution) και στη συνέχεια γίνεται σύγκριση των αποτελεσμάτων της μεθόδου με άλλες παραμετρικές μεθόδους εύρεσης της αξίας σε κίνδυνο.

4.2 ΕΚΤΙΜΗΣΗ ΤΟΥ VaR

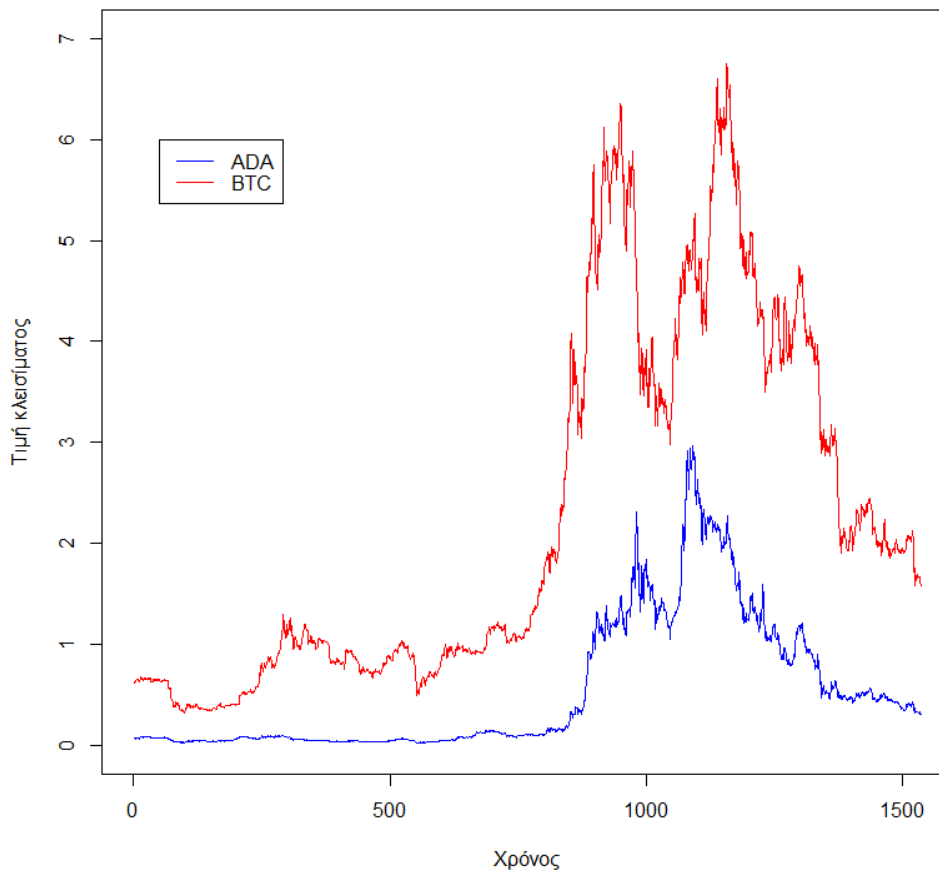
Σε ένα ρεαλιστικό οικονομικό περιβάλλον, η χάραξη στρατηγικής της διαχείρισης του επενδυτικού ρίσκου, πραγματοποιείται σε επίπεδο χαρτοφυλακίου, καθώς η λογική των επενδύσεων και η συνήθης πρακτική τους, επιτάσσει την ταυτόχρονη επένδυση σε διάφορους τίτλους. Αυτή η πρακτική έχει ως στόχο την εξομάλυνση του συνολικού κινδύνου και απωλειών με την αντιστάθμιση των όποιων αρνητικών επιδόσεων ενός τίτλου από αντίστοιχες θετικές ενός άλλου.

Συνεπώς, η γνώση των μέγιστων απωλειών ενός χαρτοφυλακίου, με βάση και μια πιθανότητα p ($VaR_p(X_1, \dots, X_n)$), κρίνεται αναγκαία. Στην περίπτωση δε, που οι κατανομές των επιμέρους τίτλων του χαρτοφυλακίου παρουσιάζουν βαριές ουρές, η εφαρμογή μοντέλων των extreme value copulas, μπορεί να εξαγάγει ασφαλέστερα αποτελέσματα σε σύγκριση με άλλες μεθόδους. Για την παρούσα εφαρμογή, χρησιμοποιήθηκαν τα κρυπτονομίσματα του Bitcoin και του Cardano. Τα δεδομένα αφορούν ημερήσιες τιμές

κλεισίματος για την περίοδο 8/9/2018 έως 30/11/2022, εξάγονται από την ιστοσελίδα Yahoo finance, μεταφορτώνονται στο στατιστικό πακέτο R, και αποτυπώνονται σε γραφήμα με βάση τον παρακάτω κώδικα:

```
> d=read.table("BTC_ADA.csv",header=TRUE,sep=";",dec=",")
> install.packages("mnormt")
> library(evd)
> library(mnormt)
> plot(d$Adj.Close.ADA,type="l",ylim=c(0,7),xlim=c(0,1537),col="blue",xlab = "Χρόνος",ylab =
"Τιμή κλεισίματος",main = "Γράφημα 4.1: Ημερήσιες τιμές των κρυπτονομισμάτων BTC και
ADA\η(8/9/2018 έως 30/11/2022)")
> lines(d$Adj.Close.BTC*0.0001,type="l",col="red")
> legend(50,6,c("ADA","BTC"),lty = c(1,1),col=c("blue","red"))
```

Γράφημα 4.1: Ημερήσιες τιμές των κρυπτονομισμάτων BTC και ADA (8/9/2018 έως 30/11/2022)



*Οι ημερήσιες τιμές κλεισίματος του BTC παρουσιάζονται σε κλίμακα $\frac{1}{10.000}$ για σκοπούς ευδιάκριτης αναπαράστασης του παραπάνω γραφήματος.

Για την εφαρμογή των μεθόδων, μας ενδιαφέρει η μελέτη των αποδόσεων των δύο παραπάνω κρυπτονομισμάτων. Ένας αρκετά δημοφιλής τρόπος υπολογισμού των αποδόσεων είναι οι λογαριθμικές αποδόσεις. Η εξίσωση η οποία χρησιμοποιείται για τον υπολογισμό των λογαριθμικών αποδόσεων είναι η ακόλουθη:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln(P_t) - \ln(P_{t-1}).$$

Στην περίπτωση όπου θέλουμε να αναλύσουμε τις αποδόσεις χρηματοοικονομικών τίτλων όπως οι μετοχές και τα κρυπτονομίσματα, συνίσταται η χρήση των λογαριθμικών και όχι των απλών αποδόσεων διότι είναι συμμετρικές, σε αντίθεση με τις απλές αποδόσεις, οι οποίες δεν είναι συμμετρικές. Επίσης οι λογαριθμικές αποδόσεις μπορεί στην απλούστερη περίπτωση να θεωρηθεί ότι ακολουθούν κανονική κατανομή (υπόδειγμα Black and Scholes). Συνήθως όμως ακολουθούν μια κατανομή με βαρύτερες ουρές, κάτι που θα ληφθεί υπόψη στη συνέχεια χρησιμοποιώντας την GEV.

Καθώς μας ενδιαφέρει η εκτίμηση των μέγιστων απωλειών ενός χαρτοφυλακίου τίτλων, η μελέτη μας επικεντρώνεται στις αρνητικές λογαριθμικές αποδόσεις. Για την πιο ομαλή, και κατανοητή αποτύπωση των αποτελεσμάτων, κάνουμε τις εξής δυο μετατροπές:

- αντιστρέφουμε το πρόσημο των λογαριθμικών αποδόσεων
- πολλαπλασιάζουμε επί 100 τις λογαριθμικές αποδόσεις

Για παράδειγμα, μια ημερήσια απόδοση της τάξεως του $-a\%$, αποτυπώνεται στα πλαίσια αυτής της εργασίας με την τιμή a . Ο κώδικας που χρησιμοποιήθηκε είναι ο ακόλουθος:

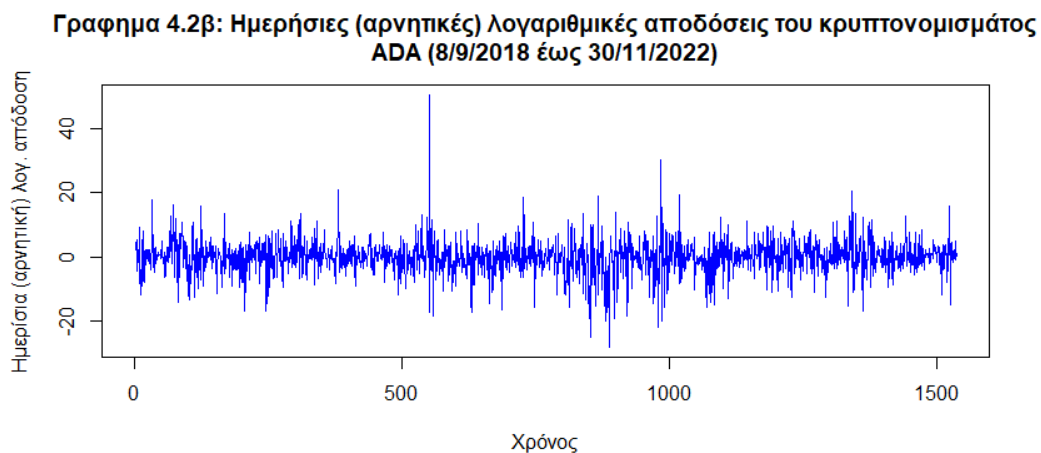
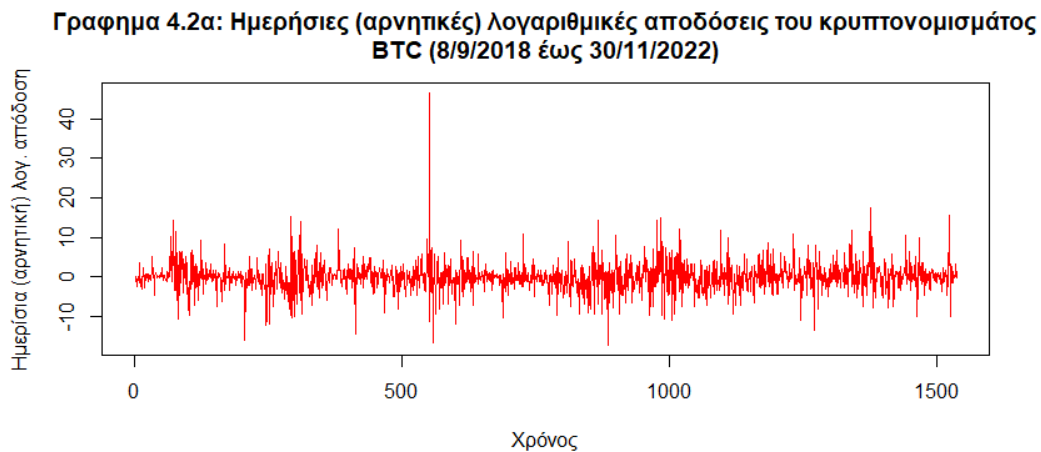
```
> Y1=log(d$Adj.Close.BTC)[1:length(d$Close)];  
> Y2=log(d$Adj.Close.ADA)[1:length(d$Close)];  
> X1=-diff(Y1);X2=-diff(Y2)
```

Στη συνέχεια, μέσω του γραφήματος των ημερήσιων (αρνητικών) λογαριθμικών αποδόσεων των κρυπτονομισμάτων του BTC και του ADA ελέγχεται η καταλληλότητα των δεδομένων της χρονοσειράς τους. Ο κώδικας που χρησιμοποιήθηκε είναι ο ακόλουθος:

```
> par(mfrow=c(2,1))
```

```
> plot(100*X1,type="l",col="red",xlab = "Χρόνος",ylab = "Ημερήσια (αρνητική) λογ. απόδοση",main = "Γράφημα 4.2α: Ημερήσιες (αρνητικές) λογαριθμικές αποδόσεις του κρυπτονομίσματος BTC\n(8/9/2018 έως 30/11/2022)");
```

```
> plot(100*X2,type="l",col="blue",xlab = "Χρόνος",ylab = "Ημερήσια (αρνητική) λογ. απόδοση",main = "Γράφημα 4.2β: Ημερήσιες (αρνητικές) λογαριθμικές αποδόσεις του κρυπτονομίσματος ADA\n(8/9/2018 έως 30/11/2022)");
```



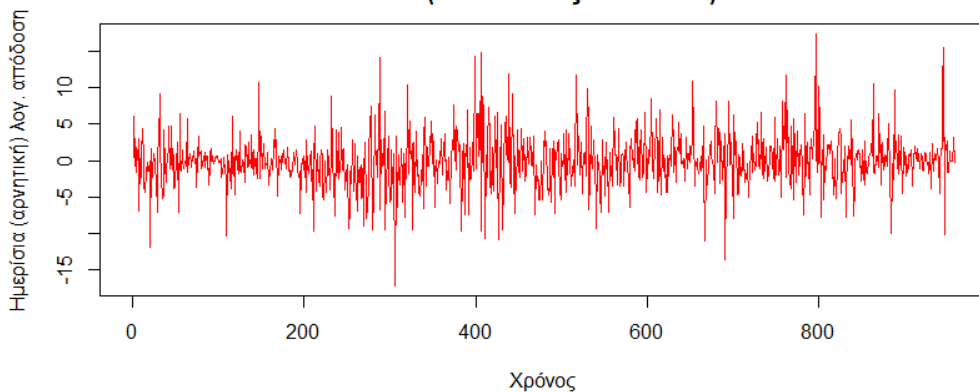
Από τα παραπάνω γραφήματα (4.2α, 4.2β), παρατηρούμε έντονη μεταβλητότητα της διασποράς των αποδόσεων και στις δύο χρονοσειρές. Πιο συγκεκριμένα παρατηρούμε ότι την περίοδο πριν την εμφάνιση της Covid-19 (400^η έως 500^η χρονική στιγμή - Q4 2019) η διακύμανση των αποδόσεων είναι αρκετά χαμηλή, ενώ μετά την εμφάνιση της Covid-19 (500^η έως 550^η χρονική στιγμή – Q1 2020) η διακύμανση των αποδόσεων είναι αρκετά υψηλή.

Δεδομένου ότι τα μοντέλα πρόβλεψης που χρησιμοποιούμε για την εκτίμηση των μέγιστων απωλειών ενός χαρτοφυλακίου τίτλων, θεωρούν σταθερή μέση τιμή και διασπορά,

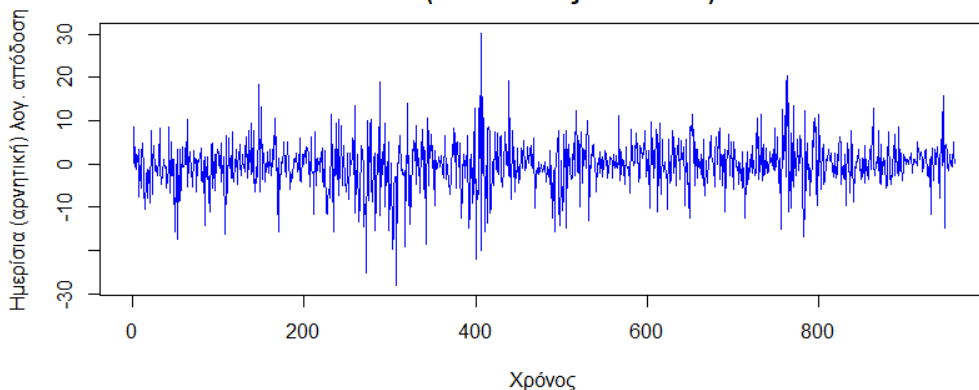
η χρήση των παραπάνω δεδομένων κρίθηκε ακατάλληλη. Για τον λόγο αυτό, επιλέχθηκε να μελετηθεί η περίοδος μετά την εμφάνιση της Covid-19 και πιο συγκεκριμένα η περίοδος 8/4/2020 έως 30/11/2022. Ο κώδικας που χρησιμοποιήθηκε είναι ο ακόλουθος:

```
> res1=100*X1[579:1536];res2=100*X2[579:1536];m=length(res1)
> plot(res1,type="l",col="red",xlab = "Χρόνος",ylab = "Ημερήσια (αρνητική) λογ. απόδοση",main = "Γράφημα 4.3α: Ημερήσιες (αρνητικές) λογαριθμικές αποδόσεις του κρυπτονομισμάτος BTC\n(8/4/2020 έως 30/11/2022)");
> plot(res2,type="l",col="blue",xlab = "Χρόνος",ylab = "Ημερήσια (αρνητική) λογ. απόδοση",main = "Γράφημα 4.3β: Ημερήσιες (αρνητικές) λογαριθμικές αποδόσεις του κρυπτονομισμάτος ADA\n(8/4/2020 έως 30/11/2022)")
```

Γραφημα 4.3α: Ημερήσιες (αρνητικές) λογαριθμικές αποδόσεις του κρυπτονομισμάτος BTC (8/4/2020 έως 30/11/2022)



Γραφημα 4.3β: Ημερήσιες (αρνητικές) λογαριθμικές αποδόσεις του κρυπτονομισμάτος ADA (8/4/2020 έως 30/11/2022)



Από τα παραπάνω γραφήματα (4.3α, 4.3β), παρατηρούμε ότι τόσο η μέση τιμή όσο και η διακύμανση δύνανται να θεωρηθούν περίπου σταθερές.

Το επόμενο βήμα της μελέτης μας, είναι να εξετάσουμε αν τα δεδομένα μας (λογαριθμικές αποδόσεις) προέρχονται από κανονικό πληθυσμό. Μια από τις κύριες επιδιώξεις της θεωρίας ακραίων τιμών είναι ο προσδιορισμός της μορφής της δεξιάς (ή αριστερής) ουράς

της κατανομής των υπό μελέτη παρατηρήσεων, ώστε να είναι δυνατή η μακροπρόθεσμη πρόβλεψη εμφάνισης ακραίων παρατηρήσεων. Συνεπώς, για να εφαρμόσουμε τη θεωρία αυτή στα δεδομένα μας, πρέπει αυτά να προέρχονται από κατανομές με βαριές δεξιές ουρές. Για τον παραπάνω έλεγχο, κατασκευάζεται για κάθε ένα κρυπτονόμισμα ιστόγραμμα συχνοτήτων και προσαρμόζεται σε αυτό τόσο η κατανομή του δείγματος όσο και κατάλληλη κανονική καμπύλη. Εκτός αυτού κατασκευάζεται κατάλληλο Q-Q διάγραμμα.

Υπενθυμίζεται ότι το Q-Q διάγραμμα είναι ένα διάγραμμα που μας επιτρέπει να εξετάσουμε αν ένα τυχαίο δείγμα X_1, X_2, \dots, X_n προέρχεται από πληθυσμό με συνάρτηση κατανομής F . Σε ένα Q-Q διάγραμμα απεικονίζονται τα n σε πλήθος σημεία:

$$(F^{-1}(p_i), x_i), \quad \text{όπου } p_i = \frac{i - a}{n + 1 - 2a} \approx \frac{i}{n} \text{ και } a = \begin{cases} \frac{3}{8}, & n \leq 10 \\ \frac{1}{2}, & n \geq 10 \end{cases}$$

(το σημείο $F^{-1}(p_i)$ αποτελεί το (κάτω) p_i ποσοστιαίο σημείο της κατανομής F).

- Αν τα σημεία του διαγράμματος είναι προσεγγιστικά κοντά στην ευθεία $y = x$ τότε δεν μπορούμε να απορρίψουμε ότι το τυχαίο δείγμα προέρχεται από πληθυσμό με συνάρτηση κατανομής $F(x)$.
- Αν τα σημεία είναι κοντά σε μια άλλη ευθεία τότε υπάρχει ένδειξη ότι το τυχαίο δείγμα προέρχεται από μια ευρύτερη οικογένεια κατανομών που τα μέλη της διαφέρουν μόνο ως προς θέση ή/και την κλίμακα με έναν εκπρόσωπο αυτής της οικογένειας να είναι η $F(x)$.
- Αν τα σημεία δεν είναι κοντά σε κάποια ευθεία και δημιουργούν μια καμπύλη προς τα άκρα της ευθείας, τότε το τυχαίο δείγμα δεν προέρχεται από πληθυσμό με συνάρτηση κατανομής ιδίου τύπου με την $F(x)$.

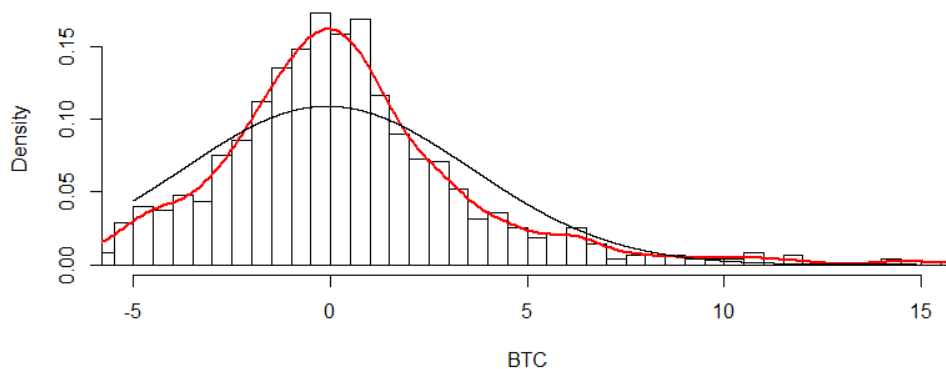
Ειδικότερα, για να ελεγχθεί αν τα δεδομένα προέρχονται από την κανονική κατανομή, τη θέση της $F(x)$ παίρνει η $\Phi(x)$ εφαρμόζοντας τη συνάρτηση *qnorm*. Όταν χρησιμοποιείται η συνάρτηση *qqnorm(x)*, η συνάρτηση *qqline(x)* προσθέτει μια ευθεία γραμμή στο κανονικό Q-Q διάγραμμα που περνά από τα σημεία $(\Phi^{-1}(0.25), x_{0.25})$ και $(\Phi^{-1}(0.75), x_{0.75})$. Ο κώδικας που χρησιμοποιήθηκε είναι ο ακόλουθος:

```
> par(mfrow=c(2,1))
```

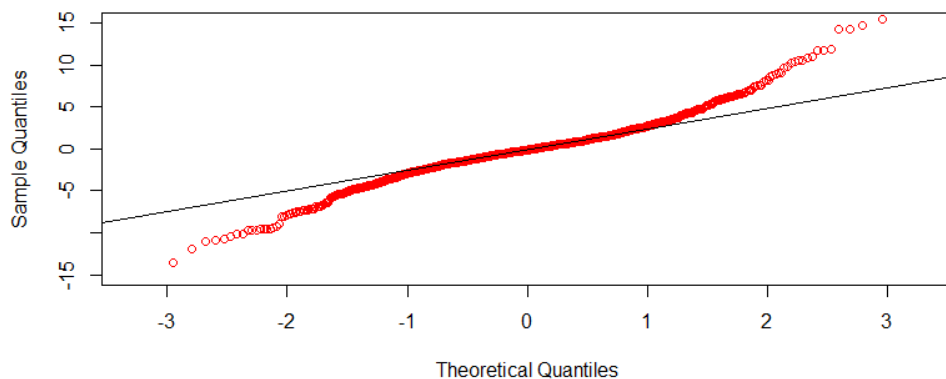
```
> hist(res1,brakes<-100,prob=TRUE,xlab="BTC",main="Γράφημα 4.4α: Κατανομή των ημερήσιων  
(αρνητικών) λογ. αποδόσεων του BTC",xlim=c(-5,15));lines(den-  
sity(res1),col="red",lwd=2);lines(seq(-5,15,0.0001),dnorm(seq(-  
5,15,0.0001),mean(res1),sd(res1)),lwd=1);qqnorm(res1,ylim=c(-15,15),col="red",main="Γράφημα  
4.4β: Normal Q-Q Plot");qqline(res1)
```

```
> hist(res2,brakes<-50,prob=TRUE,xlab="ADA",main="Γράφημα 4.5α: Κατανομή των ημερήσιων  
(αρνητικών) λογ. αποδόσεων του ADA",xlim=c(-5,20));lines(den-  
sity(res2),col="blue",lwd=2);lines(seq(-5,20,0.0001),dnorm(seq(-  
5,20,0.0001),mean(res2),sd(res2)),lwd=1);qqnorm(res2,col="blue",ylim=c(-20,20),main="Γράφημα  
4.5β: Normal Q-Q Plot");qqline(res2)
```

Γράφημα 4.4α: Κατανομή των ημερήσιων (αρνητικών) λογ. αποδόσεων του BTC

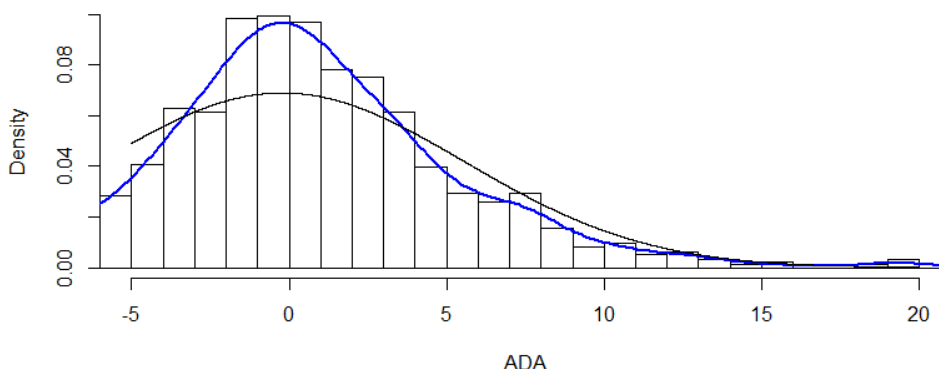


Γράφημα 4.4β: Normal Q-Q Plot

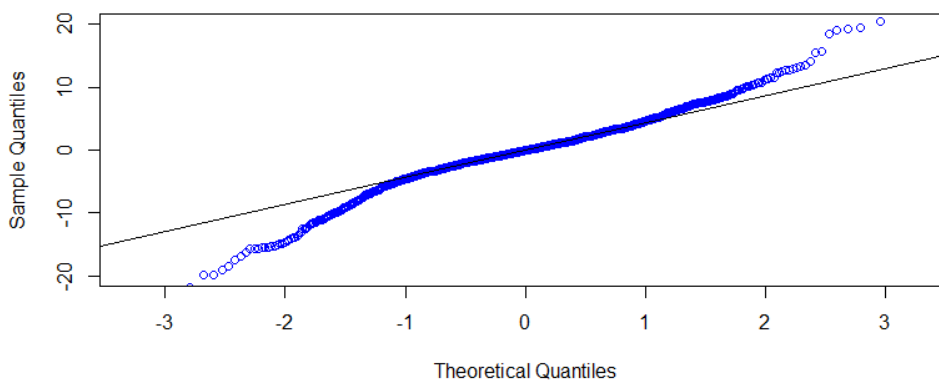


Από τα παραπάνω Γραφήματα, είναι διακριτό ότι τα δεδομένα (ημερήσιες αρνητικές λογαριθμικές αποδόσεις) από το κρυπτονόμισμα BTC απορρίπτουν την υπόθεση της κανονικότητας. Πιο συγκεκριμένα, στο γράφημα 4.4α παρατηρούμε ότι η κατανομή του δείγματος επεκτείνεται προς τη δεξιά ουρά, η οποία επίσης είναι σταθερά πιο πάνω από την εφαρμοζόμενη κανονική καμπύλη. Στο γράφημα 4.4β παρατηρούμε ότι τα σημεία δεν «αγκαλιάζουν» την ευθεία γραμμή και αποκλίνουν αισθητά στα άκρα της ευθείας.

Γράφημα 4.5α: Κατανομή των ημερίσιων (αρνητικών) λογ. αποδόσεων του ADA



Γράφημα 4.5β: Normal Q-Q Plot



Αντίστοιχες παρατηρήσεις είναι διακριτές και στα δεδομένα (ημερήσιες αρνητικές λογαριθμικές αποδόσεις) του κρυπτονομίσματος ADA. Η κατανομή του δείγματος στο διάγραμμα 4.5.α επεκτείνεται προς τη δεξιά ουρά, η οποία επίσης είναι σταθερά πιο πάνω από την εφαρμοζόμενη κανονική καμπύλη. Τέλος, η απόκλιση των σημείων από την ευθεία γραμμή υποδηλώνει την απόρριψη της κανονικότητας για την κατανομή του δείγματος.

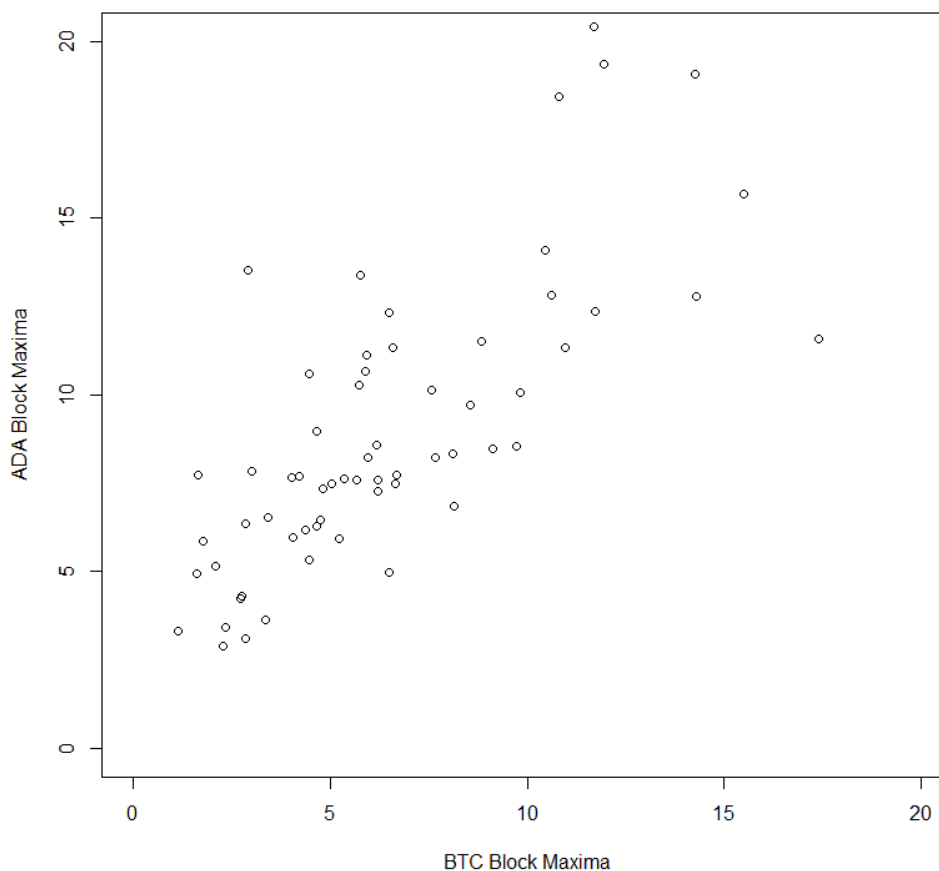
Έχοντας απορρίψει την υπόθεση της κανονικότητας, προχωράμε στον προσδιορισμό της μορφής της δεξιάς ουράς της κατανομής των υπό μελέτη παρατηρήσεων μέσω της Θεωρίας ακραίων τιμών και πιο συγκεκριμένα εφαρμόζοντας την παραμετρική μέθοδο Block Maxima (βλ. κεφάλαιο 1.2). Πιο συγκεκριμένα, χωρίζουμε τις m το πλήθος παρατηρήσεις σε n το πλήθος υποσύνολα που το καθένα περιέχει k το πλήθος παρατηρήσεις ($m = kn$), τότε οι μέγιστες παρατηρήσεις από κάθε υποσύνολο (n το πλήθος) ονομάζονται block maxima (μερικά μέγιστα) και είναι της μορφής:

$$Y_i = M_k^{(i)} = \max(X_1^{(i)}, X_2^{(i)}, \dots, X_k^{(i)}), \quad i = 1, 2, \dots, n$$

Ο κώδικας που χρησιμοποιήθηκε είναι ο ακόλουθος:

```
> par(mfrow=c(1,1))
> k=15;n=floor(m/k)
> BM1=rep(0,n);for(i in 1:n){BM1[i]=max(res1[(((i-1)*k+1):(i*k]))];}
> BM2=rep(0,n);for(i in 1:n){BM2[i]=max(res2[(((i-1)*k+1):(i*k]))];}
> bvdata=c(BM1,BM2);dim(bvdata)=c(n,2)
> plot(bvdata,xlab = "BTC Block Maxima",ylab = "ADA Block Maxima", main="Γράφημα 4.6:
Διάγραμμα διασποράς των BM του BTC σε σχέση με τα BM του ADA",ylim = c(0,20),xlim = c(0,20))
```

Γράφημα 4.6: Διάγραμμα διασποράς των BM του BTC σε σχέση με τα BM του ADA



Στόχος μας είναι να ανευρεθεί η από κοινού συνάρτηση κατανομής του διανύσματος των κανονικοποιημένων μεγίστων με τη βοήθεια της Extreme Value Copula και των περιθώριων συναρτήσεων κατανομών που όπως δείξαμε σε προηγούμενο κεφάλαιο (βλ. Κεφάλαιο 1.2) συγκλίνουν σε μια συγκεκριμένη οικογένεια κατανομών: τη γενικευμένη κατανομή ακραίων τιμών (GEV).

Ένα σημαντικό θεώρημα στη θεωρία των συναρτήσεων σύζευξης (copulas) αποτελεί το θεώρημα του Sklar (1959), (βλ. Κεφάλαιο 2.1) σύμφωνα με το οποίο η δομή της εξάρτησης μεταξύ τυχαίων μεταβλητών, μπορεί να πραγματοποιηθεί μέσω των συναρτήσεων σύζευξης (copulas).

Η κλάση των διμεταβλητών μοντέλων ακραίων τιμών περιλαμβάνει πολλές οικογένειες κατανομών που μπορούν να χρησιμοποιηθούν για τη μοντελοποίηση της δομής εξάρτησης στα δεδομένα. Υπάρχουν αρκετά διαφορετικά παραμετρικά διμεταβλητά μοντέλα ακραία τιμών για τον υπολογισμό της από κοινού συνάρτησης κατανομής των μεγίστων από ακραίες παρατηρήσεις (βλ. Κεφάλαιο 2.2).

Είδαμε στο Κεφάλαιο 2 ότι αν έχουμε τα block maxima από διμεταβλητά δεδομένα τότε αυτά θα ακολουθούν οριακά μια συνάρτηση της μορφής:

$$G(z_1, z_2) = C(G_1(z_1), G_2(z_2))$$

όπου G_1, G_2 είναι σ.κ. της GEV (με γενικά διαφορετικές παραμέτρους) και C θα είναι μια συνάρτηση σύζευξης η οποία καλείται ακραίων τιμών (extreme value copula). Μάλιστα στο κεφάλαιο αυτό είχε δοθεί η γενική μορφή που πρέπει να έχει μια συνάρτηση σύζευξης ακραίων τιμών. Μια ισοδύναμη αναπαράσταση της οριακής κατανομής των διδιάστατων block maxima είδαμε στην παράγραφο 2.2.3. ότι είναι:

$$G(z_1, z_2) = \exp \left\{ -(y_1 + y_2) A \left(\frac{y_1}{y_1 + y_2} \right) \right\}$$

όπου:

$$y_j = y_j(z_j) = \left\{ 1 + \frac{\xi_j(z_j - \mu_j)}{\sigma_j} \right\}_+^{-\frac{1}{\xi_j}}$$

για $\sigma_j > 0$ και $j = 1, 2, \dots$. Οι οριακές περιθώριες συναρτήσεις είναι GEV, και είναι της μορφής $G_j(z_j) = \exp(-y_j)$. Συνεπάγεται ότι $A(0) = A(1) = 1$ και ότι η A είναι κυρτή συνάρτηση με $\max(x, 1-x) \leq A(x) \leq 1$ για όλα τα $0 \leq x \leq 1$. Σε πλήρη ανεξαρτησία $A(1/2) = 1$ ενώ σε πλήρη εξάρτηση $A(1/2) = 0.5$.

Για τους σκοπούς αυτής της εργασίας στο εξής θα χρησιμοποιούμε την πιο διαδεδομένη extreme value copula, την Gumbel (log model) EV copula (Gumbel, 1960) που ορίζεται για τη διδιάστατη περίπτωση ως εξής (βλ. Παράγραφο 2.2.3):

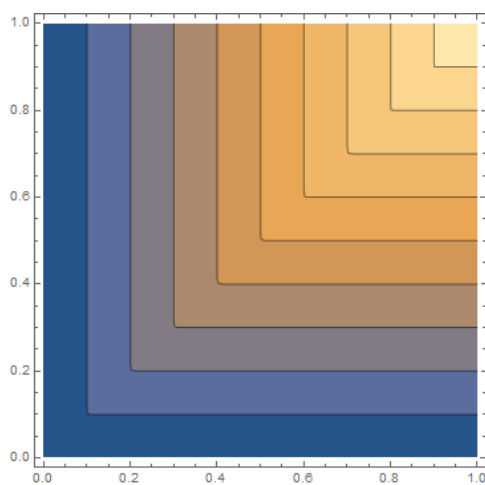
$$C(x_1, x_2) = \exp(-((- \ln x_1)^{1/r} + (- \ln x_2)^{1/r})^r)$$

και συνεπώς η οριακή κατανομή των δισδιάστατων Block maxima θα είναι:

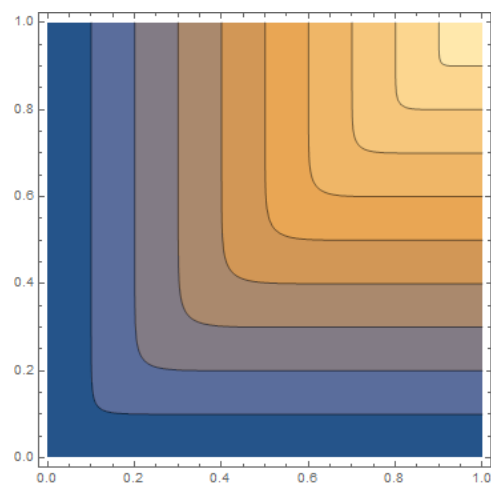
$$G(z_1, z_2) = C(G_1(z_1), G_2(z_2)) = \exp\left(-\left(y_1^{\frac{1}{r}} + y_2^{\frac{1}{r}}\right)^r\right),$$

$$y_j = \left\{1 + \frac{\xi_j(z_j - \mu_j)}{\sigma_j}\right\}_+^{\frac{1}{\xi_j}},$$

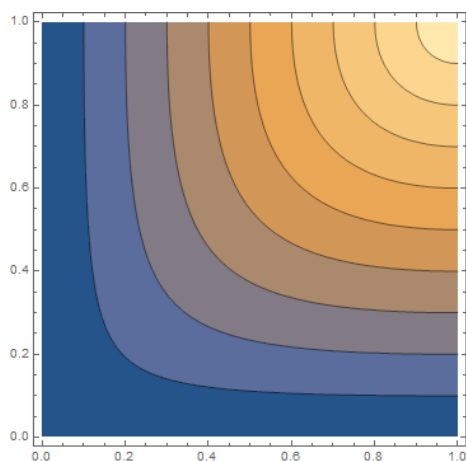
όπου με r συμβολίζεται η παράμετρος της εξάρτησης και ισχύει $0 < r \leq 1$. Πλήρης εξάρτηση επιτυγχάνεται στο όριο καθώς το r πλησιάζει το μηδέν. Ανεξαρτησία επιτυγχάνεται όταν $r = 1$. Στα παρακάτω γραφήματα δίνεται η μορφή της παραπάνω $C(x_1, x_2)$ διδιάστατης σύζευξης ακραίων τιμών (log model) για διάφορες τιμές της παραμέτρου r .



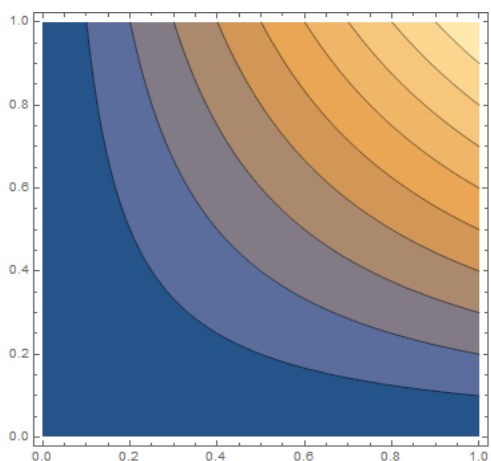
$r = 0$



$r = 0.1$



$r = 0.5$



$r = 1$

Εφαρμόζοντας το παραπάνω μοντέλο στα δεδομένα μας (B.M. του BTC , B.M. του ADA), είμαστε σε θέση να εκτιμήσουμε τόσο τις παραμέτρους των περιθώριων συναρτήσεων κατανομών, όσο και την παράμετρο της εξάρτησης. Το λογιστικό μοντέλο (log model)

μελετάται χρησιμοποιώντας το πακέτο **evd** του R (Stephenson A.). Ο κώδικας που χρησιμοποιήθηκε είναι ο ακόλουθος:

```
> fbvevd(bvdata, model = "log")> M1 <- fitted(fbvevd(bvdata, model = "log"));M1
```

Εκτίμηση των παραμέτρων των περιθώριων GEV και της εξάρτησης	
	Εκτίμηση
μ_1	4,721
σ_1	2,736
ξ_1	0,147
μ_2	6,855
σ_2	3,076
ξ_2	0,233
r	0,435

όπου:

μ_1 : η παράμετρος θέσης της κατανομής των BM του BTC

σ_1 : η παράμετρος κλίμακας της κατανομής των BM του BTC

ξ_1 : η παράμετρος σχήματος της κατανομής των BM του BTC

μ_2 : η παράμετρος θέσης της κατανομής των BM του ADA

σ_2 : η παράμετρος κλίμακας της κατανομής των BM του ADA

ξ_2 : η παράμετρος σχήματος της κατανομής των BM του ADA

r : η παράμετρος εξάρτησης των δύο μεταβλητών

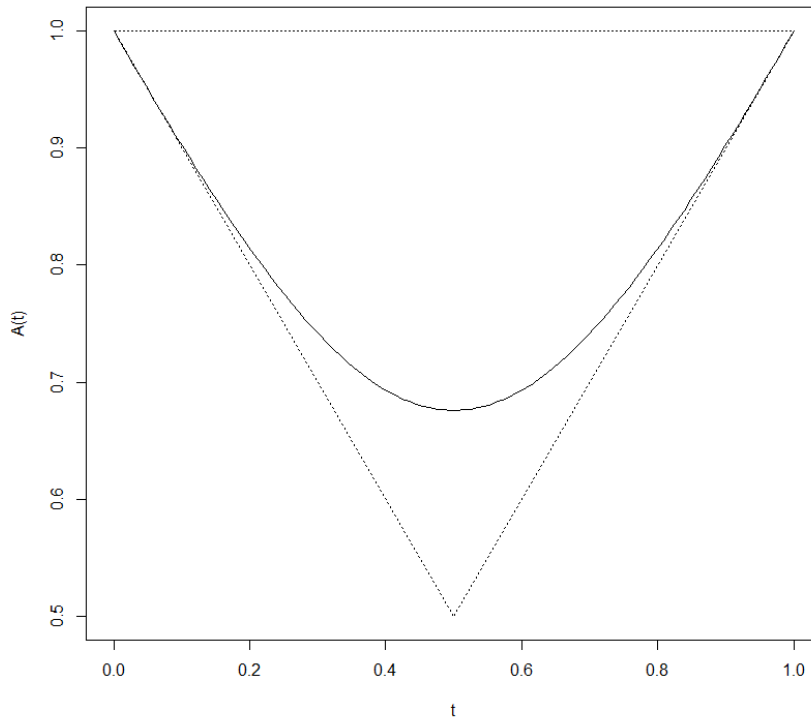
Όπως έχει ήδη αναφερθεί παραπάνω, η συνάρτηση εξάρτησης A αντιπροσωπεύει τη δομή εξάρτησης της από κοινού κατανομής και ως εκ τούτου, γνωρίζοντας την παράμετρο εξάρτησης μέσα από την εφαρμογή του παραμετρικού μοντέλου, μπορούμε να κατασκευάσουμε κατάλληλο γράφημα απεικόνισης της συνάρτησης εξάρτησης. Για το log model που εξετάζουμε είναι εύκολο να επαληθευτεί ότι:

$$A(x) = (x^{1/r} + (1 - x)^{1/r})^r, x \in [0,1].$$

Ο κώδικας που χρησιμοποιήθηκε για την απεικόνιση της συνάρτησης εξάρτησης είναι ο ακόλουθος:

```
> abvevd(dep = M1["dep"], model = "log", plot = TRUE, main="Γράφημα 4.7: Αποτύπωση της εξάρτησης μεταξύ των BM του BTC και των BM του ADA " )
```

Γράφημα 4.7: Αποτύπωση της εξάρτησης μεταξύ των BM του BTC και των BM του ADA



Έχοντας εκτιμήσει τις παραμέτρους των περιθώριων συναρτήσεων κατανομής GEV, καθώς και την παράμετρο της εξάρτησης, είμαστε σε θέση να κατασκευάσουμε τόσο την από κοινού συνάρτηση κατανομής, όσο και την από κοινού συνάρτηση πυκνότητας. Ο κώδικας που χρησιμοποιήθηκε είναι ο ακόλουθος:

```
> mu1=M1["loc1"];sigma1=M1["scale1"]; xi1=M1["shape1"]; r=M1["dep"];
> mu2=M1["loc2"];sigma2=M1["scale2"]; xi2=M1["shape2"]
> cdf1=function(x,y){pbvevd(c(x,y),mar1 = c(mu1,sigma1,xi1),mar2 = c(mu2,sigma2,xi2),
+      dep = r, model = "log",lower.tail = TRUE)}
> pdf1=function(x,y){dbvevd(c(x,y),mar1 = c(mu1,sigma1,xi1),mar2 = c(mu2,sigma2,xi2),
+      dep = r, model = "log")}
```

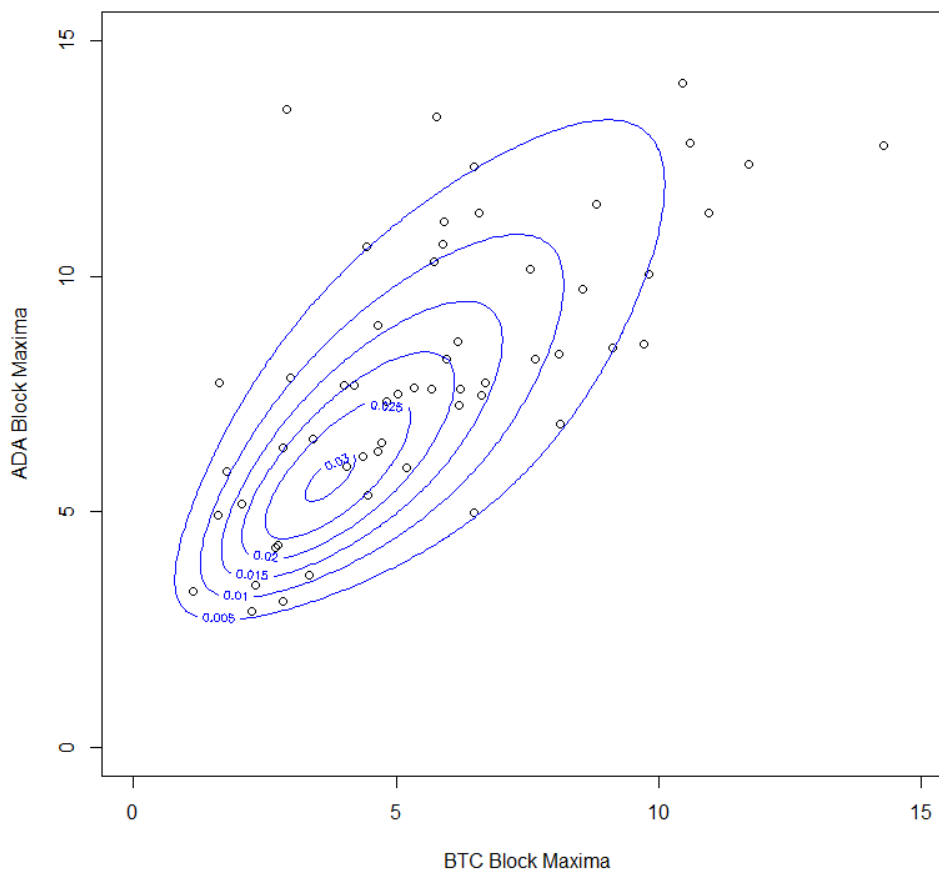
Πιο συγκεκριμένα, παρουσιάζονται τα διαγράμματα κατανομών της από κοινού συνάρτησης κατανομής και της από κοινού συνάρτησης πυκνότητας μεταξύ των BM του BTC και των BM του ADA υπό τη μορφή ισοϋψών καμπυλών (contour plots).

Τα contour plots είναι ένας τρόπος για να αποτυπωθεί μια τρισδιάστατη επιφάνεια σε ένα δισδιάστατο επίπεδο. Επί της ουσίας, είναι μια γραφική τεχνική για την αναπαράσταση μιας τρισδιάστατης επιφάνειας με την απεικόνιση σταθερών καμπυλών z (contours), σε διδιάστατη μορφή. Δηλαδή, δεδομένης μιας τιμής για το z , σχεδιάζονται

καμπύλες για τη σύνδεση των συντεταγμένων (x,y) όπου λαμβάνει χώρα αυτή η τιμή z . Ο κώδικας που χρησιμοποιήθηκε είναι ο ακόλουθος:

```
> x=seq(-3,15,0.1);nx=length(x);y=seq(-5,15,0.1);ny=length(y)#;nx;ny;nx*ny
> z=rep(0,nx*ny);dim(z)=c(nx,ny)
> for(i in 1:nx){for(j in 1:ny){z[i,j]=pdf1(x[i],y[j])}}
> plot(bvdata,xlab = "BTC Block Maxima",ylab = "ADA Block Maxima",main="Γράφημα 4.8: Από
κοινού συνάρτηση πυκνότητας πιθανότητας \nuπό τη μορφή ισοψών καμπυλών (Contour
plot)",ylim = c(0,15),xlim = c(0,15));contour(x, y, z,add=TRUE,col="blue",nlevels = 10)
> for(i in 1:nx){for(j in 1:ny){z[i,j]=cdf1(x[i],y[j])}}
> plot(bvdata,xlab = "BTC Block Maxima",ylab = "ADA Block Maxima",main="Γράφημα 4.9: Απο
κοινού συνάρτηση κατανομής \nuπο τη μορφή ισοψών καμπυλών (Contour plot)",ylim =
c(0,15),xlim = c(0,15));contour(x, y, z,add=TRUE,col="blue",nlevels = 10)
```

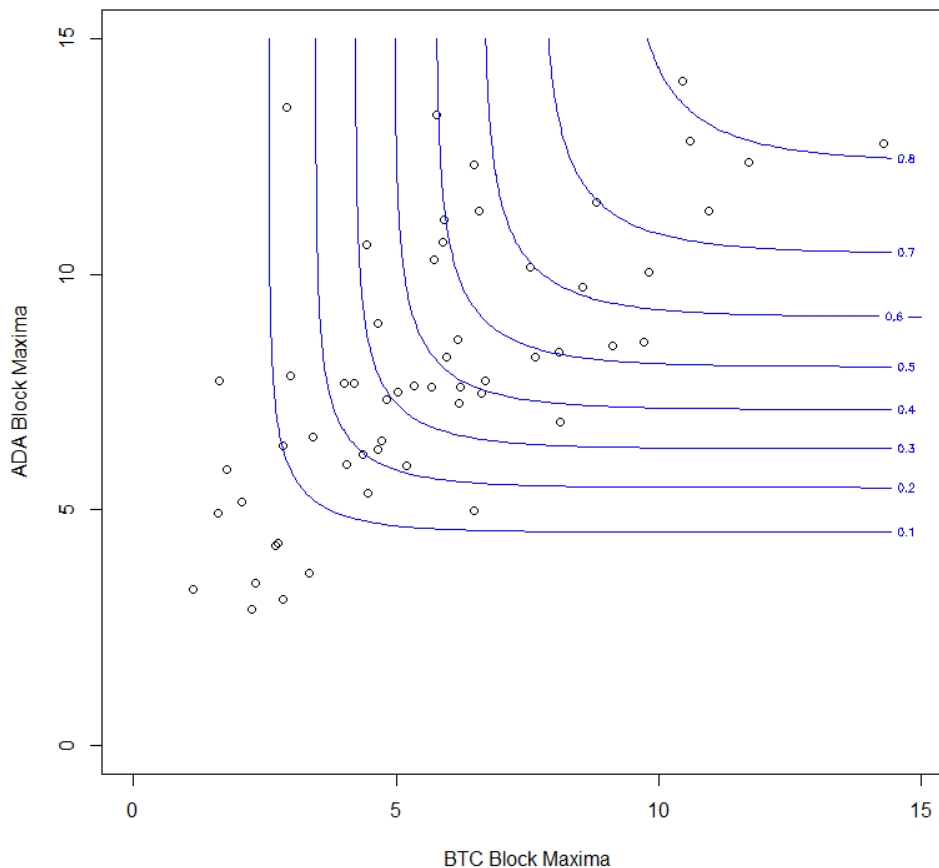
**Γράφημα 4.8: Από κοινού συνάρτηση πυκνότητας πιθανότητας
υπο την μορφή ισοψών καμπυλών (Contour plot)**



Κάθε καμπύλη στο διάγραμμα ισοψών καμπυλών αντιστοιχεί σε μια σταθερή ποσότητα της από κοινού συνάρτησης πυκνότητας. Τα διάφορα σημεία (x, y) που βρίσκονται πάνω στις ισοψείς καμπύλες, ικανοποιούν τη σταθερή τιμή της από κοινού συνάρτησης πυκνότητας. Στο συγκεκριμένο γράφημα, παρατηρούμε ότι στις περιοχές όπου η

πυκνότητα λαμβάνει όλο και μεγαλύτερες τιμές εμφανίζονται όλο και περισσότερα σημεία (δισδιάστατα Block Maxima) και συνεπώς μπορούμε να συμπεράνουμε ότι οι εκτιμήσεις των παραμέτρων από το διμεταβλητό μοντέλο ακραίων τιμών, έχουν οδηγήσει σε μια καλή προσαρμογή της κατανομής στα δεδομένα.

**Γράφημα 4.9: Απο κοινού συνάρτηση κατανομής
υπο την μορφή ισοϋψών καμπυλών (Contour plot)**



Αντίστοιχα, κάθε καμπύλη στο διάγραμμα ισοϋψών καμπυλών αντιστοιχεί σε μια σταθερή ποσότητα της από κοινού συνάρτησης κατανομής.

Η περιοχή που σχηματίζεται, προβάλλοντας οποιοδήποτε σημείο του άξονα $x'x$ ($y'y$) κάθετα έως ότου συναντήσει την ισοϋψή καμπύλη και προεκτείνοντάς το πάλι κάθετα έως ότου συναντήσει τον άξονα $y'y$ ($x'x$), έχει πιθανότητα ίση με τη σταθερή τιμή της συγκεκριμένης ισοϋψούς καμπύλης.

Γνωρίζοντας από τη μονοδιάστατη περίπτωση (βλ. Κεφάλαιο 1) ότι το VaR μιας ζημίας X με επίπεδο εμπιστοσύνης $p \in (0,1)$ είναι το κατώφλι που ξεπερνάει η ζημία X με πιθανότητα p , ισχύει ότι:

$$P(X > VaR_p(X)) = p \Leftrightarrow VaR_p(X) = F_X^{-1}(1 - p)$$

και ότι όταν πρόκειται για μονοδιάστατες BM παρατηρήσεις που ακολουθούν GEV κατανομή, τότε η εκτίμηση για το VaR προκύπτει ως το άνω p ποσοστημόριο της προσαρμοσμένης στα δεδομένα κατανομής GEV με παραμέτρους τους ε.μ.π. $\hat{\mu}, \hat{\sigma}, \hat{\xi}$ και είναι της μορφής:

$$\widehat{VaR}_p(M) = GEV^{-1}((1-p)^n; \hat{\mu}, \hat{\sigma}, \hat{\xi})$$

όπου n είναι το πλήθος των Blocks. Για τα διδιάστατα BM μπορούμε να εργαστούμε αναλόγως. Στη διδιάστατη περίπτωση αντίστοιχα θα πρέπει να βρούμε σημείο στο επίπεδο, έστω $VaR_p(X, Y) = (v_1, v_2)$, τέτοιο ώστε:

$$F_{X,Y}(v_1, v_2) = P(X \leq v_1, Y \leq v_2) = 1 - p.$$

Προφανώς όμως είναι άπειρα τα σημεία που ικανοποιούν την παραπάνω ισότητα και θα βρίσκονται πάνω σε μια καμπύλη. Έτσι, μπορούμε για κάθε δεδομένο v_1 να προσδιορίσουμε το v_2 . Αν θεωρήσουμε τα διδιάστατα BM:

$$(M_1, M_2) = (\max(X_1, \dots, X_k), \max(Y_1, \dots, Y_k))$$

Θα ισχύει σε σχέση με τα παραπάνω σημεία (v_1, v_2) ότι:

$$\begin{aligned} P(M_1 \leq v_1, M_2 \leq v_2) &= P(\max(X_1, \dots, X_k) \leq v_1, \max(Y_1, \dots, Y_k) \leq v_2) \\ &= P(X_1 \leq v_1, \dots, X_k \leq v_1, Y_1 \leq v_2, \dots, Y_k \leq v_2) = F_{X,Y}(v_1, v_2)^k = (1-p)^k \end{aligned}$$

Και επειδή τα διδιάστατα BM θα ακολουθούν προσεγγιστικά τη διδιάστατη GEV (θεωρώντας ως συνάρτηση σύζευξης το log model που αναφέραμε παραπάνω) θα είναι $G(v_1, v_2) \approx (1-p)^k$ και πιο συγκεκριμένα:

$$G(v_1, v_2) = \exp\left(-\left(\left(1 + \xi_1 \frac{v_1 - \mu_1}{\sigma_1}\right)^{-\frac{1}{r\xi_1}} + \left(1 + \xi_2 \frac{v_2 - \mu_2}{\sigma_2}\right)^{-\frac{1}{r\xi_2}}\right)^r\right) \approx (1-p)^k$$

από όπου λύνοντας ως προς v_2 (για δεδομένο δηλαδή v_1) προκύπτει ότι:

$$v_2 \approx \mu_2 + \frac{\sigma_2}{\xi_2} \left(\left((-k \ln(1-p))^{\frac{1}{r}} - \left(1 + \xi_1 \frac{v_1 - \mu_1}{\sigma_1}\right)^{-\frac{1}{r\xi_1}} \right)^{-r\xi_2} - 1 \right)$$

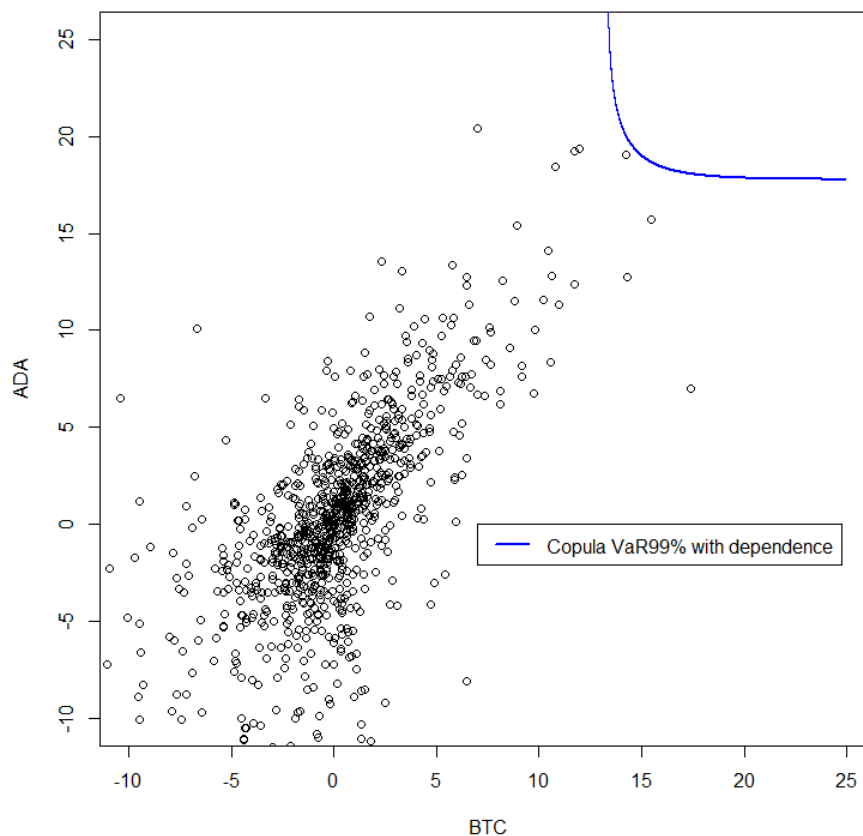
Άρα για κάθε v_1 και p βρίσκουμε από τον παραπάνω τύπο το v_2 και έτσι μπορούμε να προσδιορίσουμε την καμπύλη των σημείων $(v_1, v_2) = VaR_p(X, Y)$ τέτοια ώστε:

$$F_{X,Y}(VaR_p(X, Y)) = 1 - p.$$

Ο κώδικας που χρησιμοποιήθηκε για την απεικόνιση της παραπάνω καμπύλης του διδιάστατου VaR (μαζί με το αρχικό δείγμα) είναι ο ακόλουθος:

```
> p=0.995;
> plot(res1,res2,xlab="BTC",ylab="ADA",xlim = c(-10,25),ylim = c(-10,25),main="Γράφημα 4.10:
Διάγραμμα διασποράς των ημερήσιων (αρνητικών) λογαριθμικών αποδόσεων του BTC σε
σχέση με του ADA")
> invF2=function(v1,p){mu2+sigma2/xi2*((-k*log(p))^(1/r) -
+ (1+xi1*(v1-mu1)/sigma1)^(-1/(r*xi1)) )^(-r*xi2)-1)}
> for(v1 in seq(0,25,0.001)){points(v1,invF2(v1,p),col = "blue", cex = 0.3)}
> legend(8,0,"Copula VaR99% with dependence",lty=1,col="blue",lwd=2)
```

Γράφημα 4.10: Διάγραμμα διασποράς των ημερήσιων (αρνητικών) λογαριθμικών αποδόσεων του BTC σε σχέση με του ADA



Η καμπύλη στο παραπάνω διάγραμμα αντιστοιχεί στη σταθερή τιμή $1 - p = 0.995$ της από κοινού συνάρτησης κατανομής των (αρνητικών) λογαριθμικών αποδόσεων των δύο κρυπτονομισμάτων.

Η περιοχή που σχηματίζεται, προβάλλοντας οποιοδήποτε σημείο του άξονα $x'x$ ($y'y$) κάθετα έως ότου συναντήσει την καμπύλη και προεκτείνοντάς το πάλι κάθετα έως ότου

συναντήσει τον άξονα $y'y (x'x)$, έχει πιθανότητα ίση με τη σταθερή τιμή της συγκεκριμένης καμπύλης.

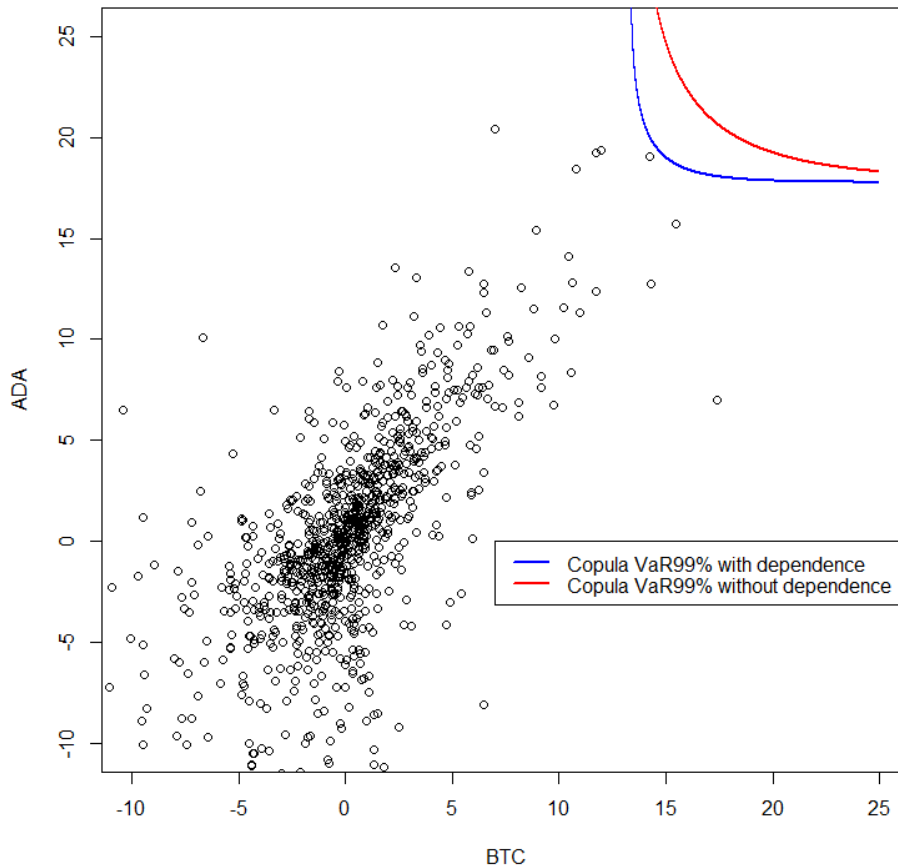
Αν θεωρούσαμε ότι δεν υπάρχει εξάρτηση στα δύο κρυπτονομίσματα του συγκεκριμένου χαρτοφυλακίου, η από κοινού οριακή συνάρτηση κατανομής θα έπαιρνε τη μορφή ($r = 1$):

$$G(z_1, z_2) = \exp(-(y_1 + y_2)) \Rightarrow G(z_1, z_2) = GEV(z_1) GEV(z_2)$$

Με τον ίδιο τρόπο όπως και πριν, εκτιμούμε τον υψηλότερο δυνατό συνδυασμό αρνητικών αποδόσεων από το χαρτοφυλάκιο τίτλων που διακρατάμε στο $(1 - p)$ των δυνατών περιπτώσεων. Ο κώδικας που χρησιμοποιήθηκε είναι ο ακόλουθος:

```
> invF3=function(v1,p){mu2+sigma2/xi2*((-k*log(p))^(1/1) -
+ (1+xi1*(v1-mu1)/sigma1)^(-1/(1*xi1)) )^(-1*xi2)-1)}
> for(v1 in seq(0,25,0.001)){points(v1,invF3(v1,p),col = "red", cex = 0.3)}
> legend(7,0,c("Copula VaR99% with dependence", "Copula VaR99% without dependence"),lty=c(1,1),col=c("blue", "red"),lwd=c(2,2))
```

Γράφημα 4.10: Διάγραμμα διασποράς των ημερίσιων (αρνητικών) λογαριθμικών αποδόσεων του BTC σε σχέση με του ADA



Τέλος, μπορούμε να εξετάσουμε την εκτίμηση του διδιάστατου VaR αν είχαμε υποθέσει απλούστερα ότι οι αποδόσεις των κρυπτονομισμάτων ακολουθούν κανονική κατανομή, με μέση τιμή μ και τυπική απόκλιση σ . Η δισδιάστατη τυχαία μεταβλητή $\mathbf{X} = (X_1, X_2)$ ακολουθεί τη διδιάστατη κανονική κατανομή με παραμέτρους $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ και ρ αν η από κοινού συνάρτηση πυκνότητας δίνεται από τον τύπο:

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2(1-\rho^2)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}T^2\right), \quad x_1, x_2 \in R$$

όπου:

$$T^2 = \frac{1}{1-\rho^2} \left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho \left(\frac{x_1-\mu_1}{\sigma_1}\right) \left(\frac{x_2-\mu_2}{\sigma_2}\right) \right]$$

Αρχικά εκτιμούμε τις παραμέτρους μ_1, μ_2 για κάθε σετ αποδόσεων των επιμέρους κρυπτονομισμάτων. Για τον υπολογισμό της τυπικής απόκλισης του χαρτοφυλακίου εκτιμάται ο 2×2 πίνακας συνδιακύμανσης:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{2,1} & \sigma_2^2 \end{bmatrix}$$

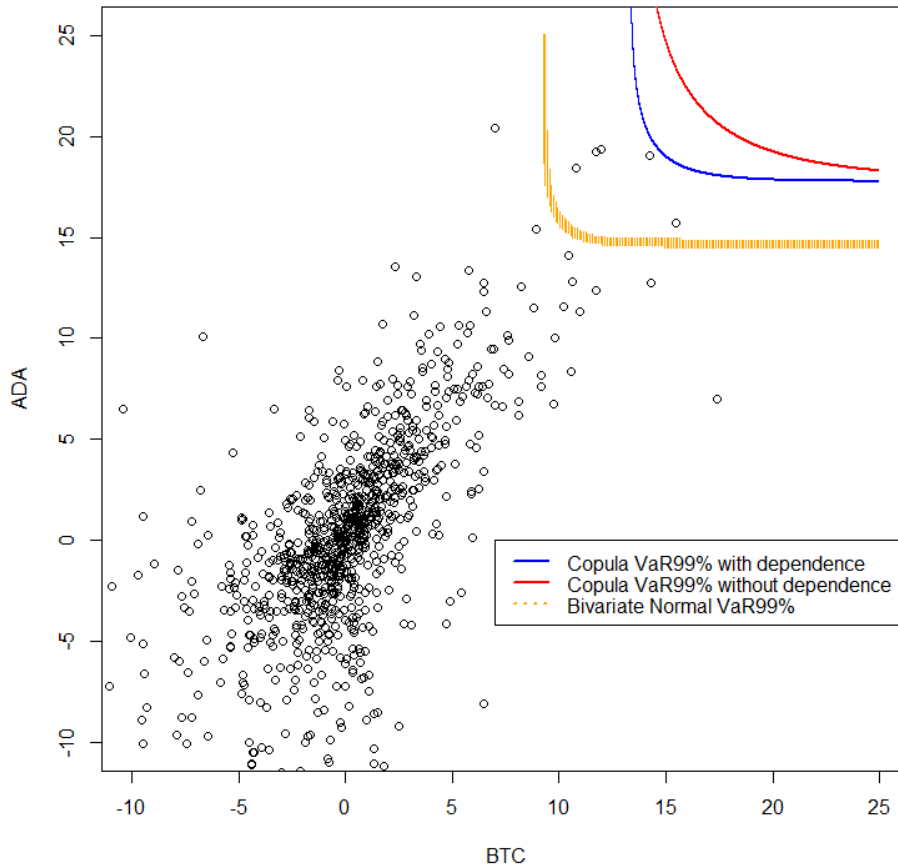
όπου τα σ_1^2 και σ_2^2 αποτελούν τις διακυμάνσεις των επιμέρους τίτλων ενώ τα $\sigma_{1,2} = \sigma_{2,1} = \rho\sigma_1\sigma_2$ αποτελούν τη συνδιακύμανση η οποία παρέχει πληροφορίες για τον τρόπο με τον οποίο οι δύο μεταβλητές σχετίζονται μεταξύ τους.

Έχοντας εκτιμήσει τις παραμέτρους, είμαστε σε θέση να κατασκευάσουμε τη διδιάστατη κανονική κατανομή και εν συνεχεία εκτιμάται η αξία σε κίνδυνο του χαρτοφυλακίου. Ο κώδικας που χρησιμοποιήθηκε είναι ο ακόλουθος (με τη βοήθεια του πακέτου **mnormt**):

```
> library(mnormt)
> x <- seq(-10, 25, 0.1); y <- seq(-10, 25, 0.1)
> mu <- c(mean(res1), mean(res2))
> sigma <- matrix(c(var(res1), cov(res1, res2), cov(res1, res2), var(res2)), nrow=2)
> f <- function(x, y) pmnorm(cbind(x, y), mu, sigma); z <- outer(x, y, f)
> x=seq(0,25,0.1);nx=length(x);y=seq(0,25,0.1);ny=length(y);z=rep(0,nx*ny);dim(z)=c(nx,ny)
> for(i in 1:nx){for(j in 1:ny)
+ {if(f(x[i],y[j])>p-0.0005 & f(x[i],y[j])<p+0.0005){points(x[i],y[j],col = "orange", cex = 0.2)}}}
```

```
> legend(7,0,c("Copula VaR99% with dependence", "Copula VaR99% without dependence", "Bivariate Normal VaR99%"),lty=c(1,1,3),col=c("blue", "red", "orange"),lwd=c(2,2))
```

Γράφημα 4.10: Διάγραμμα διασποράς των ημερίσιων (αρνητικών) λογαριθμικών αποδόσεων του BTC σε σχέση με του ADA



4.3 ΣΥΜΠΕΡΑΣΜΑΤΑ

Το παραπάνω γράφημα αποτελεί την τελική εικόνα της παρούσας μελέτης εφαρμόζοντας 3 διαφορετικά μοντέλα για την εκτίμηση του VaR99.5%. Με βάση λοιπόν την εικόνα αυτή προκύπτουν τα εξής τελικά συμπεράσματα:

Αρχικά, οι καμπύλες στο παραπάνω διάγραμμα αντιστοιχούν στη σταθερή τιμή $1 - p = 0.995$ της από κοινού συνάρτησης κατανομής. Στην πράξη μπορούμε να προσδιορίσουμε την τιμή v_1 (π.χ. προσδιορίζοντας το κεφαλαιακό απόθεμα ασφαλείας για τη μια μεταβλητή) και να βρούμε το v_2 (ώστε να βρούμε το κεφαλαιακό απόθεμα ασφαλείας για την άλλη μεταβλητή).

Το πρώτο μοντέλο εκτίμησης (μπλε γραμμή) χρησιμοποιεί την από κοινού συνάρτηση κατανομής όπως αυτή προκύπτει από την εφαρμογή των συναρτήσεων σύζευξης σε

συνδυασμό με την Πολυμεταβλητή Θεωρία Ακραίων Τιμών μέσω της συνάρτησης σύζευξης ακραίων παρατηρήσεων η οποία περιγράφει επίσης τη δομή εξάρτησης μεταξύ των δύο κρυπτονομισμάτων.

Το δεύτερο μοντέλο εκτίμησης (κόκκινη γραμμή) αποτελεί μια απλουστευμένη μορφή του πρώτου μοντέλου, όπου δεν λαμβάνεται υπόψη η εξάρτηση μεταξύ των δυο κρυπτονομισμάτων (θεωρούμε δηλαδή ότι οι δυο αυτές μεταβλητές είναι ανεξάρτητες). Παρατηρούμε ότι η καμπύλη του δεύτερου μοντέλου (κόκκινη) τοποθετείται πιο ψηλά από την καμπύλη του πρώτου μοντέλου. Ως εκ τούτου, προκύπτει ότι όταν δεν μοντελοποιείται η δομή εξάρτησης μεταξύ δυο μεταβλητών (τις θεωρούμε δηλαδή ανεξάρτητες) το μοντέλο φαίνεται να υπερεκτιμάει το VaR. Δηλαδή θα προέκυπτε ότι χρειαζόμαστε μεγαλύτερο κεφαλαιακό απόθεμα ασφαλείας από το πραγματικό (δηλαδή θα δεσμεύαμε μεγαλύτερο κεφάλαιο χωρίς να είναι απαραίτητο).

Το τρίτο μοντέλο εκτίμησης (κίτρινη γραμμή) χρησιμοποιεί την από κοινού συνάρτηση κατανομής όπως αυτή προκύπτει από την εφαρμογή της διδιάστατης κανονικής κατανομής. Παρόλο που είναι μια αρκετά απλοποιημένη μέθοδος για την κατασκευή της από κοινού συνάρτησης κατανομής και εν συνεχεία της εκτίμησης του VaR, χρησιμοποιείται κατά κόρον από χρηματοπιστωτικά ιδρύματα καθώς παρέχει ευκολία στους υπολογισμούς. Σε αυτή την περίπτωση, παρατηρούμε η καμπύλη του τρίτου μοντέλου (κίτρινη) τοποθετείται πιο χαμηλά από την καμπύλη του πρώτου μοντέλου (μπλε). Προκύπτει δηλαδή ότι για χρηματοοικονομικά δεδομένα με βαριές ουρές (όπως π.χ. αρκετές μετοχές και τα περισσότερα κρυπτονομίσματα), το διδιάστατο κανονικό μοντέλο υποεκτιμά το VaR. Αυτό συμβαίνει διότι με την κανονική δεν λαμβάνονται υπόψη οι ενδεχόμενες βαριές ουρές των δεδομένων (η κανονική έχει λεπτή ουρά) ενώ, όπως παρατηρήσαμε, το μοντέλο ακραίων τιμών δεν παρουσιάζει αυτό το πρόβλημα. Επομένως αν σε αυτή την περίπτωση εκτιμηθεί το VaR μέσω κανονικής κατανομής τότε θα παρατηρείται κάλυψη της ζημίας μέσω του κεφαλαιακού αποθέματος ασφαλείας με πιθανότητα μικρότερη από την επιθυμητή (99.5%).

Άρα τελικά η βέλτιστη μέθοδος εκτίμησης του VaR είναι αυτή που βασίζεται στη διδιάστατη Θεωρία Ακραίων Τιμών λαμβάνοντας υπόψη την εξάρτηση μεταξύ των δύο μεταβλητών ($\hat{r} = 0.435$) καθώς και τις βαριές τους ουρές ($\hat{\xi}_1 = 0.147$, $\hat{\xi}_2 = 0.233$ στην περίπτωση που εξετάσαμε).

ΒΙΒΛΙΟΓΡΑΦΙΑ

Ελληνική

Αντζουλάκος Δ., (2017), *Γλώσσα Προγραμματισμού R με Εφαρμογές στον Αναλογισμό*, Πανεπιστήμιο Πειραιώς, Πανεπιστημιακές Σημειώσεις.

Μπούτσικας Μ., (2008), *Θεωρία Ακραίων Τιμών*, Πανεπιστήμιο Πειραιώς, Πανεπιστημιακές Σημειώσεις.

Ξένα

ActEd, (2019), *Risk Modelling and Survival Analysis*, The Actuarial Education Company on behalf of the Institute and Faculty of Actuaries.

Bouyé E., (2002), *Multivariate Extremes at Work for Portfolio Risk Measurement*, Financial Econometrics Research Centre, CUBS, London & HSBC Asset Management Europe (SA), Paris.

Cherubini U. & Luciano E. & Vecchiato W., (2004), *Copula Methods in Finance*, John Wiley & Sons, Ltd.

Coles S., (2001), *An Introduction to Statistical Modeling of Extreme Values*, London: Springer.

Dupuis D. J. & Jones B. L., (2006), *Multivariate Extreme Value Theory and its Usefulness in Understanding Risk*, North American Actuarial Journal.

Fawcett L. & Walshaw D., (2008), *Modelling Environmental Extremes*, Newcastle University, Lecture Notes.

Huang W., (2014), *An Introduction to Extreme Value Analysis*, Purdue University, Lecture Notes.

Joe H., (1997) *Multivariate Models and Dependence Concepts*, London: Chapman & Hall.

Joe H. & Xu J., (1996) *The Estimation Method of Inference Functions for Margins for Multivariate Models*, Technical Report No. 166, Department of Statistics, University of British Columbia, Vancouver.

- Malevergne Y. & Sornette D., (2006), *Extreme Financial Risks from Dependence to Risk Management*, Springer.
- Reiss R.-D. & Thomas M., (2007), *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*, Birkhäuser.
- Rocco M., (2011), *Extreme Value Theory for Finance: A Survey*, Questioni di Economia e Finanza (Occasional Papers) 99, Bank of Italy, Economic Research and International Relations Area.
- Shih J.H. & Louis T.A., (1995), *Inferences on the Association Parameter in Copula Models for Bivariate Survival Data*. Biometrics.
- Stephenson A. G., (2003), *Simulating Multivariate Extreme Value Distributions of Logistic Type*, Extremes.
- Stephenson A., (2018), *Functions for Extreme Value Distributions*, R Package 'evd'.
- Stephenson A., (2018), *Statistics of Multivariate Extremes*. R-project.
- Tawn, J. A., (1988), *Bivariate Extreme Value Theory: Models and Estimation*, Biometrika.
- Yamai Y. & Yoshihara T., (2005), *Value-at-risk versus expected shortfall: A practical perspective*, Journal of Banking & Finance 29.



