

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ
ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ**

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

**ΕΚΤΙΜΗΣΗ ΤΟΥ ΚΙΝΔΥΝΟΥ ΣΕ ΕΝΑ
ΑΣΦΑΛΙΣΤΙΚΟ ΧΑΡΤΟΦΥΛΑΚΙΟ
ΕΞΑΡΤΗΜΕΝΩΝ ΚΙΝΔΥΝΩΝ ΜΕΣΩ ΤΗΣ
ΘΕΩΡΙΑΣ ΤΩΝ ΣΥΝΔΕΣΜΩΝ**

Ιωάννα Κ. Καρακωνσταντίνου

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των
απαιτήσεων για την απόκτηση του Μεταπτυχιακού Διπλώματος
Ειδίκευσης στην Αναλογιστική Επιστήμη και Διοικητική Κινδύνου

Πειραιάς
Σεπτέμβριος 2021

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ
ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ**

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

**ΕΚΤΙΜΗΣΗ ΤΟΥ ΚΙΝΔΥΝΟΥ ΣΕ ΕΝΑ
ΑΣΦΑΛΙΣΤΙΚΟ ΧΑΡΤΟΦΥΛΑΚΙΟ
ΕΞΑΡΤΗΜΕΝΩΝ ΚΙΝΔΥΝΩΝ ΜΕΣΩ ΤΗΣ
ΘΕΩΡΙΑΣ ΤΩΝ ΣΥΝΔΕΣΜΩΝ**

Ιωάννα Κ. Καρακωνσταντίνου

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των
απαιτήσεων για την απόκτηση του Μεταπτυχιακού Διπλώματος
Ειδίκευσης στην Αναλογιστική Επιστήμη και Διοικητική Κινδύνου

Πειραιάς
Σεπτέμβριος 2021

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

Τα μέλη της Επιτροπής ήταν:

- Αναπληρωτής Καθηγητής: Μπούτσικας Μιχαήλ (Επιβλέπων)
- Αναπληρωτής Καθηγητής: Ψαρράκος Γεώργιος
- Επίκουρος Καθηγητής: Πιτσέλης Γεώργιος

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

UNIVERSITY OF PIRAEUS



**DEPARTMENT OF STATISTICS
AND INSURANCE SCIENCE**

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

**RISK ESTIMATION OF AN INSURANCE
PORTFOLIO OF DEPENDENT RISKS
USING COPULAS**

By

Ioanna K. Karakonstantinou

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

September 2021

Στην οικογένεια μου

Ευχαριστίες

Ολοκληρώνοντας την διπλωματική μου εργασία θα ήθελα να ευχαριστήσω αρχικά τον Καθηγητή του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς και επιβλέποντα, κ. Μπούτσικα Μιχαήλ για την πολύτιμη βοήθεια και καθοδήγηση του που συνέβαλαν καθοριστικά στην εκπόνηση της διπλωματικής μου εργασίας. Επιπλέον, θα ήθελα να ευχαριστήσω τους καθηγητές του Μεταπτυχιακού Προγράμματος Σπουδών Αναλογιστικής Επιστήμης και Διοικητικής Κινδύνου για τις γνώσεις που μου μετέφεραν. Τέλος, θα ήθελα να ευχαριστήσω την οικογένεια μου και τους φίλους μου για την κατανόηση, την υπομονή, την συμπαράσταση και την στήριξη που μου προσέφεραν και συνεχίζουν να προσφέρουν σε κάθε βήμα της ζωής μου.

Περίληψη

Η διαχείριση των κινδύνων αποτελεί καθημερινή πρόκληση για τις ασφαλιστικές εταιρείες και τα χρηματοπιστωτικά ιδρύματα και βασικό βήμα στην ανάλυση αυτών είναι η κατανόηση της εξάρτησης μεταξύ των αποτελεσμάτων. Η παρούσα εργασία επικεντρώνεται στις συναρτήσεις σύζευξης (Copulas), ένα εργαλείο για την κατανόηση των σχέσεων μεταξύ των τυχαίων μεταβλητών. Στο πρώτο κεφάλαιο εξετάζουμε μέτρα κινδύνου τα οποία χρησιμοποιούνται ευρέως στην ασφάλιση και τα χρηματοοικονομικά, παρουσιάζουμε ιδιότητες αυτών και μεθόδους εκτίμησης τους. Στην συνέχεια, περιγράφουμε το θεώρημα Sklar, θεμέλιο στην θεωρία των συζεύξεων, παρουσιάζουμε διδιάστατες συναρτήσεις σύζευξης και τις ιδιότητες αυτών και κλείνουμε το δεύτερο κεφάλαιο με την Αρχιμήδεια οικογένεια συζεύξεων. Τέλος, μελετάμε ένα ασφαλιστικό χαρτοφυλάκιο Αστικής Ευθύνης αυτοκινήτου, προσαρμόζουμε μέσω της γλώσσας προγραμματισμού R κατάλληλη σύζευξη στα δεδομένα μας και με βάση αυτήν υπολογίζουμε την αντίστοιχη δεσμευμένη Αξία σε κίνδυνο.

Abstract

Risk management is a daily challenge for insurance companies and financial institutions. A substantial step in risk analysis is the understanding of the dependence structure between one dimensional results. The current dissertation focuses on Copulas, a tool for understanding the relationship between random variables. In the first chapter, we present risk measures which are widely used in insurance and finance along with their properties and methods for their assessment. In the second chapter we describe Sklar's theorem which is a fundamental result of the theory of Copulas. We focus on bivariate Copulas and their properties and at the end of the chapter we present the class of Archimedean Copulas. In the final 3rd chapter we study a non-life insurance portfolio of Motor third part liability losses. We numerically investigate the class of optimally fitted copula on our bivariate data through R statistical package. Finally, based on the most appropriate copula, we calculate portfolio's conditional Value at Risk.

Περιεχόμενα

| | |
|---|----|
| Εισαγωγή | 18 |
| Μέτρα Κινδύνου | 20 |
| 1.1 Μέτρα Κινδύνου | 21 |
| 1.1.1 Διασπορά | 21 |
| 1.1.2 Αξία σε κίνδυνο (Value at Risk)..... | 22 |
| 1.1.3 Αναμενόμενο έλλειμμα (Expected Shortfall)..... | 26 |
| 1.2 Εκτίμηση των μέτρων κινδύνου VaR και ES..... | 29 |
| 1.2.1 Παραμετρική Εκτίμηση..... | 29 |
| 1.2.2 Εμπειρική εκτίμηση από ιστορικά δεδομένα | 30 |
| 1.2.3 Μέθοδος Variance-Covariance..... | 32 |
| Συναρτήσεις Σύζευξης | 36 |
| 2.1 Συναρτήσεις Σύζευξης (Copulas)..... | 37 |
| 2.1.1 Το Θεώρημα Sklar..... | 38 |
| 2.1.2 Σύνδεση Copula Συναρτήσεων και Συναρτήσεων Πυκνότητας Πιθανότητας..... | 40 |
| 2.2 Οικογένειες Διδιάστατων συναρτήσεων σύζευξης | 42 |
| 2.3 Ιδιότητες των συναρτήσεων σύζευξης (Copulas) | 47 |
| 2.3.1 Σύζευξη επιβίωσης (Survival Copula) | 48 |
| 2.3.2 Dual και co-copulas | 49 |
| 2.3.3 Αναλλοίωτοι μονότονοι μετασχηματισμοί (Functional Invariance) | 50 |
| 2.3.4 Εξάρτηση Ουράς | 51 |
| 2.4 Αρχιμήδεια συνάρτηση σύζευξης (Archimedean Family of Copulas)..... | 52 |
| 2.4.1 Σύνδεση Αρχιμήδειας Copula και Συναρτήσεων Πυκνότητας Πιθανότητας | 53 |
| 2.4.2 Ιδιότητες της Αρχιμήδειας σύζευξης | 55 |
| Μελέτη Ασφαλιστικού χαρτοφυλακίου | 57 |
| 3.1 Εφαρμογή σε ασφαλιστικά δεδομένα ζημιών | 57 |
| 3.1.1 Δεδομένα | 57 |
| 3.1.2 Περιθώριες κατανομές των μεταβλητών BI και PD..... | 60 |
| 3.1.3 Προσαρμογή Σύζευξης σε διδιάστατα δεδομένα..... | 63 |
| 3.1.4 Προσδιορισμός της Διδιάστατης κατανομής με την χρήση Σύζευξης | 71 |
| 3.1.5 Υπολογισμός του Δεσμευμένου VaR 99.5% (Conditional VaR 99.5%) | 76 |
| ΒΙΒΛΙΟΓΡΑΦΙΑ | 80 |

Εισαγωγή

Ο κίνδυνος συνδέεται με την πιθανότητα να συμβεί κάτι επιβλαβές και συνεπάγεται την μελλοντική αβεβαιότητα σχετικά με την απόκλιση από το αναμενόμενο αποτέλεσμα. Βασικό βήμα στην διαχείριση και την ανάλυση του κινδύνου αποτελεί η δημιουργία πολύπλοκων στατιστικών μοντέλων στα οποία η κατανόηση των σχέσεων εξάρτησης μεταξύ των μονοδιάστατων αποτελεσμάτων κρίνεται αναγκαία για την αποφυγή υποεκτίμησης του κινδύνου. Σκοπός της παρούσας διπλωματικής εργασίας είναι η μελέτη μοντελοποίησης της εξάρτησης πολυδιάστατων αποτελεσμάτων μέσω της έννοιας των συναρτήσεων Σύζευξης (Copulas). Η μαθηματική έκφραση των συναρτήσεων σύζευξης παρουσιάστηκε για πρώτη φορά το 1959 από τον Sklar και η διαθέσιμη βιβλιογραφία σχετικά με τις στατιστικές ιδιότητες και τις εφαρμογές των συζεύξεων αναπτύσσεται ραγδαία τα τελευταία χρόνια. Στο πρώτο κεφάλαιο της εργασίας μελετάμε δύο μέτρα κινδύνου τα οποία χρησιμοποιούνται ευρέως στις ασφαλιστικές και τα χρηματοπιστωτικά ιδρύματα, την Αξία σε κίνδυνο (Value at Risk) και το Αναμενόμενο Έλλειμμα (Expected shortfall) και παρουσιάσουμε τις ιδιότητες και μεθόδους εκτίμησης τους. Συνεχίσουμε στο δεύτερο κεφάλαιο με την αποτύπωση του θεωρήματος Sklar, βασικό θεώρημα στην θεωρία των συζεύξεων και θεμέλιο σε πολλές εφαρμογές, καθώς διευκρινίζει τον ρόλο των συναρτήσεων σύζευξης στην σχέση που υπάρχει μεταξύ των πολυδιάστατων συναρτήσεων κατανομής και των μονοδιάστατων περιθωρίων συναρτήσεών τους. Μελετάμε διδιάστατες συναρτήσεις σύζευξης και τις ιδιότητες αυτών και κλείνουμε το δεύτερο κεφάλαιο με την Αρχιμήδεια σύζευξη. Η Αρχιμήδεια σύζευξη αποτελεί σημαντική οικογένεια συναρτήσεων, παρέχει μια σειρά μοντέλων τα οποία έχουν χρησιμοποιηθεί με επιτυχία στα πλαίσια μοντελοποίησης δεδομένων, ενώ στην Αναλογιστική επιστήμη οι περισσότερες εφαρμογές των συναρτήσεων σύζευξης επικεντρώθηκαν στην συγκεκριμένη οικογένεια. Στο τρίτο κεφάλαιο παρουσιάζουμε την μελέτη ενός ασφαλιστικού χαρτοφυλακίου Αστικής Ευθύνης αυτοκινήτου και οι περιθώριες μεταβλητές που εξετάζουμε είναι το κόστος των σωματικών βλαβών και το κόστος των υλικών ζημιών που έχουν προκληθεί από ένα ατύχημα. Προσαρμόζουμε μέσω της γλώσσας προγραμματισμού R κατάλληλη σύζευξη στα διδιάστατα δεδομένα μας και μέσω αυτής προσδιορίζουμε την διδιάστατη από κοινού κατανομή των περιθωρίων. Τέλος με βάση την κατάλληλη συνάρτησης σύζευξης που έχουμε προσδιορίσει, υπολογίζουμε και απεικονίζουμε την δεσμευμένη Αξία σε Κίνδυνο στο 99,5%.

Κεφάλαιο 1

Μέτρα Κινδύνου

Οι άνθρωποι επιζητούν καθημερινά την ασφάλεια. Η αίσθηση ασφάλειας δημιουργεί στο άτομο την εντύπωση ότι μπορεί να ικανοποιήσει τις βασικές ανάγκες του (φαγητό, καταφύγιο, ιατρική περίθαλψη κ.α.) στο παρόν και στο μέλλον. Αντίθετα, το αίσθημα ανασφάλειας καλλιεργεί αβεβαιότητα απέναντι στην έκθεση στον κίνδυνο. Κίνδυνος είναι η πιθανότητα να συμβεί κάτι επιβλαβές ή απρόσμενο και συνεπάγεται μελλοντική αβεβαιότητα σχετικά με την απόκλιση από το αναμενόμενο αποτέλεσμα. Ποικίλουν οι προτεινόμενοι ορισμοί, όμως για την κοινή κατανόηση της έννοιας σε διαφορετικές εφαρμογές και με βάση τον Διεθνή Οργανισμό Τυποποίησης (International Organization for Standardization (ISO)), ο κίνδυνος συνδέεται με την «επίδραση της αβεβαιότητας στους στόχους».

Για αιώνες, η λειτουργία των ασφαλιστών και των αντασφαλιστών ήταν η παροχή κάλυψης κινδύνου. Τις τελευταίες δεκαετίες σε αυτή την δραστηριότητα συμμετέχουν επίσης οι τράπεζες και τα χρηματοπιστωτικά ιδρύματα. Στις μέρες μας, και οι δύο ομάδες αντιμετωπίζουν την ίδια πρόκληση της συλλογής και διαχείρισης κινδύνων αναζητώντας αγορές όπου οι κίνδυνοι μπορούν να αντισταθμιστούν ή να αποδεσμευτούν.

Στην περίπτωση που η αντιστάθμιση του κινδύνου δεν είναι εφικτή, ένα κατάλληλο μέτρο κινδύνου πρέπει να καθοριστεί και, εάν χρειάζεται, να υπολογιστούν επαρκή αποθέματα. Σήμερα, το ενδιαφέρον των εταιρειών προσελκύεται από πρόσφατα ανεπτυγμένες τεχνικές μέτρησης του κινδύνου, ενώ κατά καιρούς, πολυάριθμα μέτρα κινδύνου έχουν αναπτυχθεί στην ασφάλιση και τα χρηματοοικονομικά και κυμαίνονται από τα πιο στοιχειώδη στα πιο περίτεχνα.

Η μέτρηση του κινδύνου βασίζεται σε αξιωματική προσέγγιση. Η αναζήτηση «σωστού» μέτρου είναι αβάσιμη, δεδομένου ότι κατάλληλα μέτρα κινδύνου είναι όσα συμμορφώνονται με τα αξιώματα. Επιπλέον, διαφορετικές κατηγορίες μέτρων κινδύνου αντιπροσωπεύουν διαφορετικούς τρόπους σκέψης και αντίληψης. Επομένως, στην πράξη η επιλογή ενός κατάλληλου μέτρου ποσοτικοποίησης του κινδύνου των χαρτοφυλακίων ασφάλισης εξακολουθεί να αποτελεί αντικείμενο συζήτησης.

Η χρήση των μέτρων κινδύνου ικανοποιεί διάφορους σκοπούς. Ένας εκ των πιο σημαντικών είναι ο προσδιορισμός του επιχειρηματικού κεφαλαίου και του κεφαλαίου επάρκειας. Συγκεκριμένα, μία από τις βασικές λειτουργίες της διαχείρισης κινδύνων είναι ο προσδιορισμός του κεφαλαίου που οφείλει να διατηρήσει μία ασφαλιστική εταιρεία ως περιθώριο έναντι απροσδόκητων

μελλοντικών ζημιών στο χαρτοφυλάκιο του για να ικανοποιήσει μια ρυθμιστική αρχή, η οποία ασχολείται με τη φερεγγυότητα του ιδρύματος. Επιπλέον, τα μέτρα κινδύνου χρησιμοποιούνται συχνά από τη διοίκηση ως εργαλείο για τον περιορισμό του κινδύνου που μπορεί να αναλάβει μια μονάδα μέσα στην εταιρεία. Επιπρόσθετα, τα ασφάλιστρα μιας ασφαλιστικής εταιρείας αντισταθμίζουν τον κίνδυνο των ασφαλισμένων απαιτήσεων που έχει αναλάβει. Συνεπώς, το μέγεθος αυτής της αποζημίωσης μπορεί να θεωρηθεί μέτρο του κινδύνου των αξιώσεων αυτών.

Τα μέτρα κινδύνου μπορούν να ομαδοποιηθούν σε τέσσερις διαφορετικές κατηγορίες: την προσέγγισή πλασματικού ποσού, τα μετρά ευαισθησίας παραγόντων, τα μέτρα κινδύνου βασισμένα σε κατανομές απώλειας και τα μέτρα κινδύνου βασισμένα σε υποθέσεις. Στο παρόν κεφάλαιο παρουσιάζονται μέτρα βασισμένα σε κατανομές απώλειας (loss distribution) τα οποία χρησιμοποιούνται ευρέως στην ασφάλιση και τα χρηματοοικονομικά. Συγκεκριμένα, περιγράφονται η διασπορά, η Αξία σε κίνδυνο (Value at risk), το Αναμενόμενο έλλειμμα (Expected shortfall) και οι ιδιότητες αυτών.

Τα περισσότερα σύγχρονα μέτρα του κινδύνου ενός χαρτοφυλακίου είναι στατιστικές ποσότητες που περιγράφουν την κατανομή απώλειας του χαρτοφυλακίου σε κάποιο προκαθορισμένο ορίζοντα. Οι απώλειες είναι το κεντρικό αντικείμενο ενδιαφέροντος για τη διαχείριση κινδύνων, επομένως είναι φυσικό ένα μέτρο κινδύνου να βασίζεται στη κατανομή τους. Επιπλέον, οι κατανομές ζημιών μπορούν να συγκριθούν μεταξύ χαρτοφυλακίων. Ωστόσο, η εκτίμηση της κατανομής απωλειών βασίζεται σε ιστορικά δεδομένα τα οποία λόγω μεταβαλλόμενου περιβάλλοντος είναι περιορισμένης χρήσης για την πρόβλεψη μελλοντικού κινδύνου. Ακόμη και σε σταθερό περιβάλλον, είναι δύσκολο να εκτιμηθεί με ακρίβεια η κατανομή απωλειών, ιδιαίτερα για μεγάλα χαρτοφυλάκια, και εξελιγμένα συστήματα διαχείρισης κινδύνου που βασίζονται σε περίπλοκα στατιστικά μοντέλα για τον προσδιορισμό της κατανομής ζημιών.

Κατά την έκταση του κεφαλαίου παρουσιάζονται γνωστά αποτελέσματα τα οποία αποδεικνύουμε για λόγους πληρότητας.

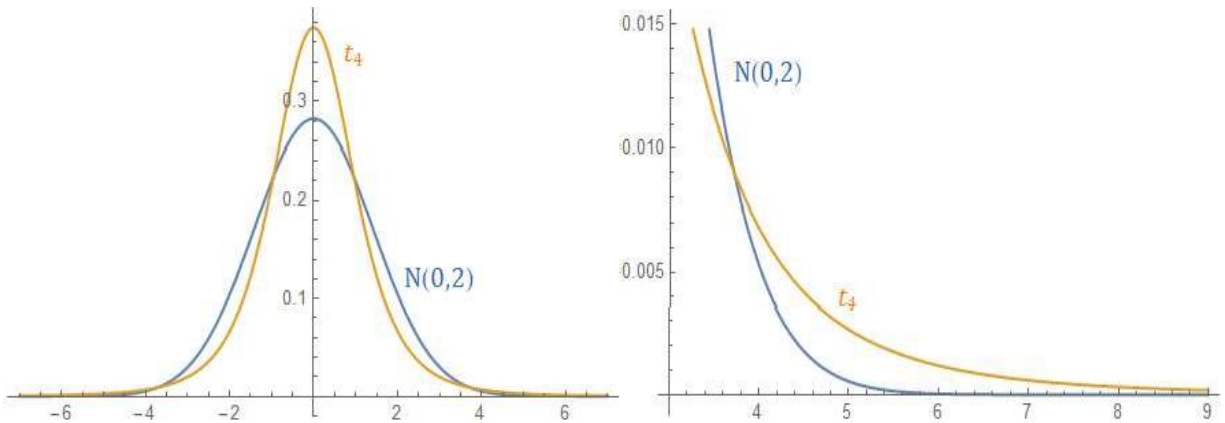
1.1 Μέτρα Κινδύνου

1.1.1 Διασπορά

Η διασπορά της κατανομής ζημιών χρησιμοποιείται συχνά ως μέτρο κινδύνου, και κυρίως στην θεωρία χαρτοφυλακίων. Αποτελεί εύληπτο και εύχρηστο μέτρο κινδύνου, ωστόσο ορίζεται μόνο για κατανομές με πεπερασμένη διασπορά. Οι ακραίες τιμές (ακραία κέρδη και ακραίες ζημιές)

επιδρούν το ίδιο στην διασπορά, ενώ δεν παρέχει καμία πληροφορία σχετικά με πιθανές μεγάλες ζημιές που μπορεί να συμβούν· το τελευταίο επιβεβαιώνεται στο παρακάτω παράδειγμα.

Έστω δύο κατανομές ζημιών $L_1 \sim N(0, \sigma^2 = 2)$ και $L_2 \sim t_4$ με την ίδια διασπορά και τις παρακάτω συναρτήσεις πυκνότητας. Παρατηρείται ότι η t_4 , η οποία χαρακτηρίζεται από βαρύτερες ουρές, έχει μεγαλύτερη πιθανότητα εμφάνισης μεγάλων ζημιών σε σύγκριση με την κανονική κατανομή.



(Τα σχήματα προέρχονται από το σύγγραμμα “Εισαγωγή στην διαχείριση Κινδύνων” Μπούτσικας Μιχαήλ (2020))

1.1.2 Αξία σε κίνδυνο (Value at Risk)

Η Αξία σε κίνδυνο (VaR, Value at Risk) είναι από τα πιο διαδεδομένα μέτρα κινδύνου στα χρηματοπιστωτικά ιδρύματα και στις ασφαλιστικές εταιρείες όπου υπολογίζεται στο πλαίσιο κεφαλαιακής επάρκειας της Φερεγγυότητας II.

Στην πράξη ο όρος VaR χρησιμοποιείται ποικιλοτρόπως. Σύμφωνα με την πιο “κοινή” έννοια, ο όρος Αξία σε κίνδυνο αναφέρεται σε ποσοστημόριο της κατανομής ζημιών. Επίσης, συχνά συναντάται η έννοια της VaR διαδικασίας η οποία αναφέρεται σε μια στατιστική προσέγγιση για την εκτίμηση ενός μοντέλου κατανομής ζημιών. Η διαδικασία VaR θα μπορούσε επίσης να χρησιμοποιηθεί για την εκτίμηση ενός άλλου μέτρου κινδύνου βάσει της κατανομής ζημιών. Επιπλέον, ο όρος «προσέγγιση VaR στη διαχείριση κινδύνων» χρησιμοποιείται συχνά και συνήθως αναφέρεται στον τρόπο με τον οποίο τα στοιχεία VaR χρησιμοποιούνται στην καθοδήγηση και την επιτήρηση μιας εταιρείας.

Ορισμός: Η Αξία σε κίνδυνο (Value at Risk) μιας ζημίας L σε επίπεδο εμπιστοσύνης $\alpha \in (0, 1)$ ισούται με

$$\begin{aligned}
\text{VaR}_\alpha(L) &= \inf \{x \in \mathbb{R}: P(L > x) \leq 1 - \alpha\} \\
&= \inf \{x \in \mathbb{R}: 1 - F_L(x) \leq 1 - \alpha\} \\
&= \inf \{x \in \mathbb{R}: F_L(x) \geq \alpha\}
\end{aligned}$$

όπου $F_L(x): \mathbb{R} \rightarrow (0, 1)$ η συνάρτηση κατανομής της ζημιάς L

Ο παραπάνω ορισμός ισχύει για διακριτές, συνεχείς και μεικτές τ.μ..

Επομένως, το VaR εκφράζει την μικρότερη τιμή x τέτοια ώστε η πιθανότητα η ζημία L να υπερβαίνει το x δεν είναι μεγαλύτερη από $(1 - \alpha)$. Εξ ορισμού το VaR σε επίπεδο εμπιστοσύνης α δεν παρέχει πληροφορίες σχετικά με τη σφοδρότητα των ζημιών που συμβαίνουν με πιθανότητα μικρότερη από $1 - \alpha$, ενώ οι συνήθεις τιμές που παίρνει το επίπεδο εμπιστοσύνης α είναι $\alpha = 0,95$ ή $\alpha = 0,99$.

Για τις συνεχείς κατανομές ζημιών (με αντιστρέψιμη συνάρτηση κατανομής) ισχύει ότι:

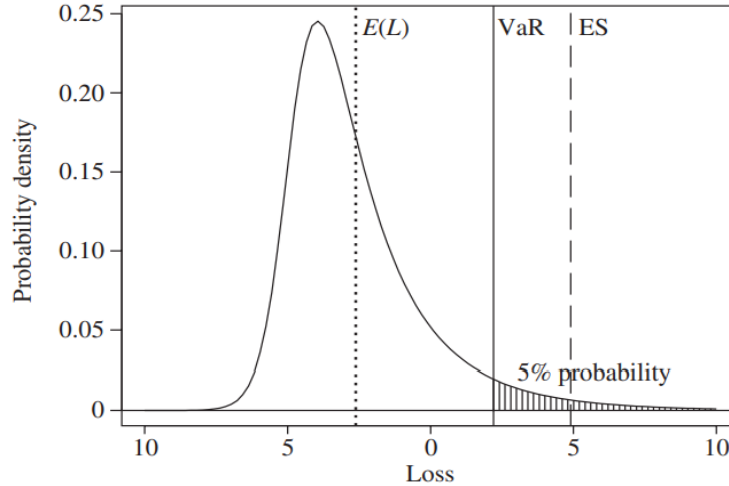
$$\text{VaR}_\alpha(L) = \inf \{x \in \mathbb{R}: F_L(x) \geq \alpha\} = \inf \{x \in \mathbb{R}: x \geq F_L^{-1}(\alpha)\} = F_L^{-1}(\alpha)$$

Ισοδύναμα

$$P(L \leq \text{VaR}_\alpha(L)) = P(L \leq F_L^{-1}(\alpha)) = F_L(F_L^{-1}(\alpha)) = \alpha$$

όπου $F_L^{-1}(\alpha)$ η αντίστροφη συνάρτηση κατανομής της L στο α .

Στο παρακάτω γράφημα απεικονίζεται η σ.π.π. της κατανομής ζημιών με το 95% VaR να επισημαίνεται ως κατακόρυφη γραμμή. Η μέση ζημιά $E(L)$ εμφανίζεται με μια διακεκομμένη γραμμή και ένα εναλλακτικό μέτρο κινδύνου γνωστό ως το 95% Αναμενόμενο έλλειμμα (Expected shortfall) επισημαίνεται με διακεκομμένη γραμμή.



(Το σχήμα προέρχεται από το σύγγραμμα “Quantitative Risk Management” των Alexander J. McNeil, Rudiger Frey και Paul Embrechts (2005))

Είναι γνωστό ότι ισχύει το ακόλουθο αποτέλεσμα.

Πρόταση: Για κάθε τ.μ. L με $\alpha \in (0,1)$, $c > 0$, $d \in \mathbb{R}$, ισχύει ότι $\text{VaR}_\alpha(cL + d) = c\text{VaR}_\alpha(L) + d$.

Απόδειξη: Από τον ορισμό του VaR έχουμε

$$\text{VaR}_\alpha(cL + d) = \inf \{x \in \mathbb{R}: P(cL + d \leq x) \geq \alpha\} = \inf \{x \in \mathbb{R}: P(L \leq (x - d)/c) \geq \alpha\}$$

Θέτουμε $x' = (x - d)/c$

$$\begin{aligned} \text{VaR}_\alpha(cL + d) &= \inf \{(cx' + d) \in \mathbb{R}: P(L \leq x') \geq \alpha\} \\ &= c \inf \{x' \in \mathbb{R}: P(L \leq x') \geq \alpha\} + d \\ &= c \text{VaR}_\alpha(L) + d \end{aligned}$$

Επομένως, ο κίνδυνος c μετοχών χαρτοφυλακίου (ίδιας AE), ο οποίος μετράται σε VaR, είναι c φορές ο κίνδυνος ενός μεριδίου αυτού του χαρτοφυλακίου. Επιπλέον, η προσθήκη ($d > 0$) ή η απόσυρση ($d < 0$) d ποσού από το χαρτοφυλάκιο μεταβάλλει τον κίνδυνο κατά το ποσό αυτό (d).

Παράδειγμα (VaR για κανονική κατανομή ζημίας):

Έστω ζημία L με συνάρτηση κατανομής F_L και $L \sim N(\mu, \sigma^2)$. Αν το επίπεδο εμπιστοσύνης $\alpha \in (0, 1)$, τότε ισχύει

$$\text{VaR}_\alpha(L) = \mu + \sigma \Phi^{-1}(\alpha)$$

όπου Φ η συνάρτηση πυκνότητας της τυποποιημένης κανονικής κατανομής.

Πράγματι, από τον ορισμό του VaR και δεδομένου ότι η σ.κ. της κανονικής κατανομής είναι αντιστρέψιμη, έχουμε

$$\begin{aligned} P(L \leq \text{VaR}_\alpha(L)) = \alpha &\Leftrightarrow P\left(\frac{L - \mu}{\sigma} \leq \frac{\text{VaR}_\alpha(L) - \mu}{\sigma}\right) = \alpha \Leftrightarrow \Phi\left(\frac{\text{VaR}_\alpha(L) - \mu}{\sigma}\right) = \alpha \\ &\Leftrightarrow \text{VaR}_\alpha(L) = \mu + \sigma\Phi^{-1}(\alpha) \end{aligned}$$

Όταν εργαζόμαστε με μέτρα κινδύνου βασισμένα σε κατανομή ζημίας, πρέπει να επιλέξουμε έναν κατάλληλο ορίζοντα και στην περίπτωση του VaR χρειάζεται να αποφασίσουμε και το επίπεδο εμπιστοσύνης α . Ο ορίζοντας διαχείρισης κινδύνου πρέπει να αντικατοπτρίζει τη χρονική περίοδο κατά την οποία η ασφαλιστική εταιρεία ή το χρηματοπιστωτικό ίδρυμα δεσμεύεται να διατηρήσει το χαρτοφυλάκιό του. Αυτή η περίοδος επηρεάζεται και εξαρτάται από νομικούς περιορισμούς και θέματα ρευστότητας και συνήθως διαφέρει μεταξύ των αγορών. Κατά την επιλογή του χρονικού ορίζοντα διαχείρισης κινδύνων της επιχείρησης, η ασφαλιστική εταιρεία ή το χρηματοπιστωτικό ίδρυμα δεν έχει άλλη επιλογή από το να χρησιμοποιήσει τον κατάλληλο ορίζοντα με βάση την αγορά στην οποία βρίσκονται οι βασικές επιχειρηματικές δραστηριότητες. Για παράδειγμα, οι ασφαλιστικές εταιρείες είναι συνήθως υποχρεωμένες να διατηρούν το χαρτοφυλάκιό τους για ένα έτος. Κατά την περίοδο αυτή δεν μπορούν να μεταβάλλουν το χαρτοφυλάκιο κατά κάποιο σημαντικό ποσό, ούτε να επαναδιαπραγματευθούν τα ασφάλιστρα τα οποία λαμβάνουν. Επιπλέον, για λόγους κεφαλαιακής επάρκειας, απαιτείται ένα υψηλό επίπεδο εμπιστοσύνης προκειμένου να υπάρχει επαρκές περιθώριο ασφαλείας. Για παράδειγμα σύμφωνα με την Φερεγγυότητα II, η Κεφαλαιακή Απαίτηση Φερεγγυότητας αντιστοιχεί στην Αξία σε Κίνδυνο (Value at Risk) με επίπεδο εμπιστοσύνης 99,5% για περίοδο ενός έτους.

Σε αρκετές περιπτώσεις υπολογισμού του VaR είναι χρήσιμη η ακόλουθη γνωστή πρόταση.

Πρόταση: Αν g είναι μια γνήσια αύξουσα συνάρτηση και $Y, g(Y)$ συνεχείς τ.μ. τότε ισχύει

$$\text{VaR}_\alpha(g(Y)) = g(\text{VaR}_\alpha(Y))$$

Αν g είναι γνήσια φθίνουσα συνάρτηση τότε ισχύει αντίστοιχα $\text{VaR}_\alpha(g(Y)) = g(\text{VaR}_{1-\alpha}(Y))$.

Απόδειξη: Εφόσον οι τ.μ. $Y, g(Y)$ είναι συνεχείς, με βάση τον ορισμό του VaR ισχύουν οι σχέσεις

$$P(Y \leq \text{VaR}_\alpha(Y)) = \alpha \quad (1)$$

$$P(g(Y) \leq \text{VaR}_\alpha(g(Y))) = \alpha \quad (2)$$

Δεδομένου ότι η συνάρτηση g είναι γνησίως αύξουσα, από την σχέση (1) προκύπτει

$$P(g(Y) \leq g(\text{VaR}_\alpha(Y))) = \alpha$$

Συγκρίνοντας με την σχέση (2) προκύπτει το ζητούμενο $\text{VaR}_\alpha(g(Y)) = g(\text{VaR}_\alpha(Y))$.

Η απόδειξη στην περίπτωση γνησίως φθίνουσας συνάρτησης g είναι ανάλογη.

1.1.3 Αναμενόμενο έλλειμμα (Expected Shortfall)

Αν και η Αξία σε κίνδυνο αποτελεί ένα ευρέως γνωστό και εύχρηστο μέτρο κινδύνου, έχει αρκετούς περιορισμούς. Για παράδειγμα, δεν παρέχει πληροφορίες σχετικά με το ύψος το οποίο μπορεί να φτάσει η ζημιά δεδομένου ότι υπερβαίνει το Value at Risk (δηλαδή σε περίπτωση που $L \geq \text{VaR}_\alpha(L)$). Το μέτρο κινδύνου Αναμενόμενο έλλειμμα (Expected shortfall) σχετίζεται στενά με την Αξία σε κίνδυνο (VaR) και δεδομένου των ανεπαρκειών του τελευταίου, προτιμάται συχνά κατά την διαχείριση κινδύνου.

Ορισμός: Το Αναμενόμενο Έλλειμμα (Expected Shortfall) μια ζημιάς L με συνεχή συνάρτηση κατανομής F_L και επίπεδο εμπιστοσύνης $\alpha \in (0,1)$ ορίζεται ως η δεσμευμένη μέση τιμή

$$\text{ES}_\alpha(L) = E(L \mid L \geq \text{VaR}_\alpha(L))$$

Η παραπάνω σχέση αναλύεται και ως εξής θέτοντας $v = \text{VaR}_\alpha(L)$

$$\text{ES}_\alpha(L) = E(L \mid L \geq v) = \frac{E(L I_{[L \geq v]})}{P(L \geq v)} = \frac{1}{1 - \alpha} \int_v^\infty x f_L(x) dx$$

όπου f_L σ.π.π. της συνεχούς τ.μ. L . Θέτουμε $y = F_L(x)$, $x = F_L^{-1}(y)$, $\frac{dy}{dx} = \frac{dF_L(x)}{dx} = f_L(x)$ και επομένως, έχουμε

$$\int_v^\infty x f_L(x) dx = \int_{F_L(v)}^{F_L(\infty)} F_L^{-1}(y) dy = \int_a^1 \text{VaR}_y(L) dy$$

Συνεπώς, το Αναμενόμενο έλλειμμα μπορεί να ερμηνευθεί ως η αναμενόμενη ζημία δεδομένου ότι υπερβαίνει το Var (δηλαδή ανήκει στο χειρότερο $100(1 - \alpha)\%$ των περιπτώσεων).

Στα πλαίσια της Βασιλείας III, δρομολογήθηκε η αντικατάσταση της Αξίας σε κίνδυνο (VaR) για τη μέτρηση του κινδύνου αγοράς από το Αναμενόμενο έλλειμμα (ES).

Για διακριτές κατανομές έχουν προταθεί αρκετοί ορισμοί για το Αναμενόμενο έλλειμμα, ωστόσο δεν θα επεκταθούμε καθώς στην συνέχεια του κεφαλαίου θα εξετάσουμε μόνο την περίπτωση των συνεχών τ.μ.. Το ακόλουθο γνωστό αποτέλεσμα είναι χρήσιμο για τον υπολογισμό του ES .

Πρόταση: Αν L συνεχής τ.μ. τότε $\text{ES}_\alpha(cL + d) = c\text{ES}_\alpha(L) + d$, $d \in \mathbb{R}$, $c > 0$.

Απόδειξη: Με βάση τον ορισμό του ES , ισχύει

$$\begin{aligned} \text{ES}_\alpha(cL + d) &= E(cL + d \mid cL + d \geq \text{VaR}_\alpha(cL + d)) \\ &= E(cL + d \mid cL + d \geq c\text{VaR}_\alpha(L) + d) \\ &= E(cL + d \mid L \geq \text{VaR}_\alpha(L)) \\ &= cE(L \mid L \geq \text{VaR}_\alpha(L)) + d = c\text{ES}_\alpha(L) + d \end{aligned}$$

Παρακάτω παρουσιάζουμε παραδείγματα υπολογισμού του Αναμενόμενου ελλείμματος για δύο γνωστές συνεχείς κατανομές (π.χ. βλ. Μπούτσικας (2020)).

Παράδειγμα: Αν L τ.μ. η οποία ακολουθεί κανονική κατανομή με μέση τιμή μ και διασπορά σ^2 ($L \sim N(\mu, \sigma^2)$) και $\alpha \in (0, 1)$ τότε

$$\text{ES}_\alpha(L) = \mu + \sigma \frac{\varphi(\Phi^{-1}(\alpha))}{1 - \alpha},$$

όπου φ και Φ η σ.π.π. και η σ.κ. της τυποποιημένης κανονικής κατανομής αντίστοιχα. Πράγματι, γνωρίζουμε από προηγούμενη εφαρμογή ότι $\text{VaR}_\alpha(L) = \mu + \sigma\Phi^{-1}(\alpha)$ όταν $L \sim N(\mu, \sigma^2)$. Έστω $L' \sim N(0,1)$, τότε $\text{VaR}_\alpha(L') = \Phi^{-1}(\alpha)$

$$\begin{aligned} \text{ES}_\alpha(L') &= \frac{1}{1 - \alpha} \int_{\Phi^{-1}(\alpha)}^{\infty} x\varphi(x) dx \\ &= \frac{1}{1 - \alpha} \int_{\Phi^{-1}(\alpha)}^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{1-\alpha} \left[-\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right]_{\Phi^{-1}(\alpha)}^{\infty} \\
&= \frac{1}{1-\alpha} \frac{1}{\sqrt{2\pi}} e^{-\frac{\Phi^{-1}(\alpha)^2}{2}} = \frac{\varphi(\Phi^{-1}(\alpha))}{1-\alpha}
\end{aligned}$$

Επομένως, αν $L = \mu + \sigma L' \sim N(\mu, \sigma^2)$ τότε

$$ES_{\alpha}(L) = ES_{\alpha}(\mu + \sigma L') = \mu + \sigma ES_{\alpha}(L') = \mu + \sigma \frac{\varphi(\Phi^{-1}(\alpha))}{1-\alpha}.$$

Παράδειγμα: Έστω $L \sim$ Εκθετική (λ) με παράμετρο λ .

Γνωρίζουμε ότι για μια μη αρνητική τ.μ. X με συνάρτηση κατανομής F και συνάρτηση επιβίωσης $\bar{F} = 1 - F$ ισχύει $E(X) = \int_0^{\infty} P(X > x) dx$. Επομένως,

$$\begin{aligned}
E(X | X > v) &= \int_0^{\infty} P(X > x | X > v) dx \\
&= \int_0^v P(X > x | X > v) dx + \int_v^{\infty} P(X > x | X > v) dx \\
&= \int_0^v 1 dx + \int_0^{\infty} P(X > v + y | X > v) dy \\
&= v + \int_0^{\infty} \frac{\bar{F}(v+y)}{\bar{F}(v)} dy
\end{aligned}$$

Με την χρήση της παραπάνω σχέσης και εφαρμόζοντας όπου $v = \text{VaR}_{\alpha}(L)$, έχουμε

$$\begin{aligned}
ES_{\alpha}(L) &= E(L | L > v) \\
&= v + \frac{1}{1-\alpha} \int_0^{\infty} \bar{F}(v+y) dy \\
&= v + \frac{1}{1-\alpha} \int_0^{\infty} e^{-\lambda(v+y)} dy \\
&= v + \frac{1}{1-\alpha} \frac{1}{\lambda} e^{-\lambda v} \quad (1)
\end{aligned}$$

Από τον ορισμό του VaR για τ.μ. $Y \sim$ Εκθετική (λ) ισχύει

$$\begin{aligned}
P(Y > \text{VaR}_{\alpha}(Y)) &= 1 - \alpha \Leftrightarrow 1 - F_Y(\text{VaR}_{\alpha}(Y)) = 1 - \alpha \\
&\Leftrightarrow e^{-\lambda \text{VaR}_{\alpha}(Y)} = 1 - \alpha
\end{aligned}$$

$$\Leftrightarrow \text{VaR}_\alpha(Y) = -\frac{1}{\lambda} \ln(1 - \alpha)$$

Επομένως, η σχέση (1) γίνεται $\text{ES}_\alpha(L) = \text{VaR}_\alpha(L) + \frac{1}{\lambda}$

Για περισσότερες πληροφορίες στα Μέτρα κινδύνου ο ενδιαφερόμενος αναγνώστης μπορεί να ανατρέξει στα συγγράμματα M. Denuit, J. Dhaene, M. Goovaerts, R. Kaas (2005), Alexander J. McNeil, Rudiger Frey, Paul Embrechts (2005) και M. Μπούτσικας (2020).

1.2 Εκτίμηση των μέτρων κινδύνου VaR και ES

Στην πράξη δεν είναι συνήθως γνωστή η κατανομή που ακολουθεί μια ζημιά L ενός χαρτοφυλακίου. Συνεπώς, θα πρέπει να εκτιμήσουμε τα μέτρα κινδύνου βασιζόμενοι σε ιστορικά στοιχεία. Κατά την απλούστερη περίπτωση, θεωρούμε ένα τυχαίο δείγμα ιστορικών ζημιών ενός χαρτοφυλακίου $L_1, L_2, L_3, \dots, L_n$ που αντιστοιχούν σε προηγούμενες χρονικές περιόδους, ακολουθούν την ίδια κατανομή και είναι ανεξάρτητες.

1.2.1 Παραμετρική Εκτίμηση

Αρχικά θα εξετάσουμε την εκτίμηση των VaR και ES μέσω ενός παραδείγματος, όπου θεωρούμε ότι οι τ.μ. L_i ακολουθούν γνωστή κατανομή με άγνωστες παραμέτρους.

Έστω $L_i \sim N(\mu, \sigma^2)$ όπου μ, σ^2 άγνωστες παράμετροι. Όπως έχουμε αποδείξει σε προηγούμενες εφαρμογές, ισχύουν οι παρακάτω σχέσεις:

$$\text{VaR}_\alpha(L) = \mu + \sigma \Phi^{-1}(\alpha), \quad \text{ES}_\alpha(L) = \mu + \sigma \frac{\varphi(\Phi^{-1}(\alpha))}{1 - \alpha}$$

για να μπορέσουμε να εκτιμήσουμε τα παραπάνω μέτρα κινδύνου μέσω των

$$\widehat{\text{VaR}}_\alpha(L) = \hat{\mu} + \hat{\sigma} \Phi^{-1}(\alpha), \quad \widehat{\text{ES}}_\alpha(L) = \hat{\mu} + \hat{\sigma} \frac{\varphi(\Phi^{-1}(\alpha))}{1 - \alpha},$$

θα πρέπει αρχικά να εκτιμήσουμε τις άγνωστες παραμέτρους, χρησιμοποιώντας για παράδειγμα τις γνωστές εκτιμήτριες του μέσου και της διασποράς

$$\hat{\mu} = \bar{L} = \frac{1}{n} \sum_{i=1}^n L_i, \quad \hat{\sigma}^2 = S_L^2 = \frac{1}{n-1} \sum_{i=1}^n (L_i - \bar{L})^2$$

Η παραμετρική εκτίμηση επιτυγχάνει ικανοποιητικές εκτιμήσεις των μέτρων κινδύνου, όμως ενέχει ο κίνδυνος εφαρμογής λανθασμένου παραμετρικού μοντέλου, για παράδειγμα, οι ζημιές να είναι από μια κατανομή με διαφορετικά χαρακτηριστικά.

Άρα, εκφράζουμε τα μέτρα κινδύνου συναρτήσει των άγνωστων παραμέτρων, για παράδειγμα $VaR_\alpha(L) = v(\theta)$, $ES_\alpha(L) = s(\theta)$, και στην συνέχεια εκτιμώντας τις παραμέτρους (συνήθως μέσω της μεθόδου της μέγιστης πιθανοφάνειας), λαμβάνουμε τα εκτιμώμενα μέτρα κινδύνου

$$\widehat{VaR}_\alpha(L) = v(\widehat{\theta}), \quad \widehat{ES}_\alpha(L) = s(\widehat{\theta})$$

Όπου η εκτιμήτρια μέγιστης πιθανοφάνειας (ε.μ.π.) $\widehat{\theta} = (\theta_1, \theta_2, \dots, \theta_k)$ του θ , ισούται με την τιμή του θ η οποία μεγιστοποιεί την συνάρτηση του λογαρίθμου της πιθανοφάνειας (log-likelihood)

$$\ell(\theta) = \log \prod_{i=1}^n f(L_i; \theta) = \sum_{i=1}^n \log f(L_i; \theta).$$

1.2.2 Εμπειρική εκτίμηση από ιστορικά δεδομένα

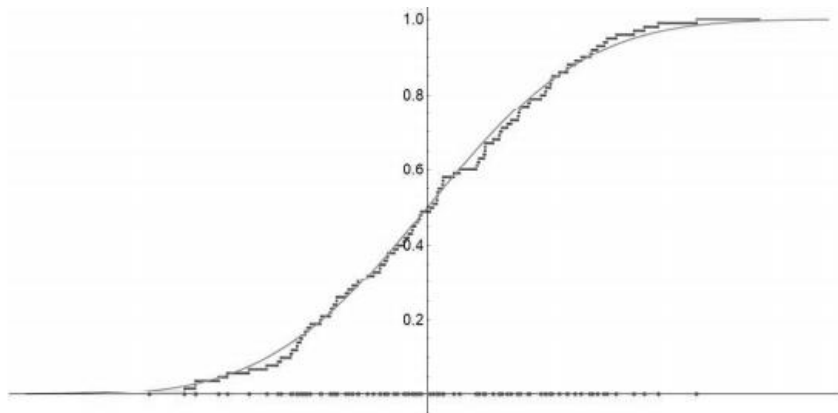
Είδαμε την εκτίμηση των μέτρων κινδύνου VaR και ES μέσω της εφαρμογής ενός παραμετρικού μοντέλου. Εναλλακτικά, μπορούμε να βασιστούμε στην εμπειρική κατανομή ιστορικών δεδομένων θεωρώντας ανεξάρτητες και ισόνομες τ.μ. $L_1, L_2, L_3, \dots, L_n$ (i.i.d.) οι οποίες δεν ακολουθούν κάποια συγκεκριμένη κατανομή.

Μέσω της εμπειρικής συνάρτησης κατανομής εκτιμάται η άγνωστη σ.κ F των L_i . Δηλαδή,

$$\widehat{F}_n(x) = \widehat{P}(L \leq x) = \frac{\text{πλήθος } \{L_i : L_i \leq x\}}{n} = \frac{1}{n} \sum_{i=1}^n I_{[L_i \leq x]}$$

όπου $I_{[L_i \leq x]}$ η δείκτρια συνάρτηση $I_{[L_i \leq x]} = \begin{cases} 1, & L_i \leq x \\ 0, & L_i > x \end{cases}$

Και από τον Νόμο των Μεγάλων Αριθμών ισχύει ότι $\widehat{F}_n(x) \xrightarrow{n \rightarrow \infty} E(I_{[L_i \leq x]}) = F(x)$

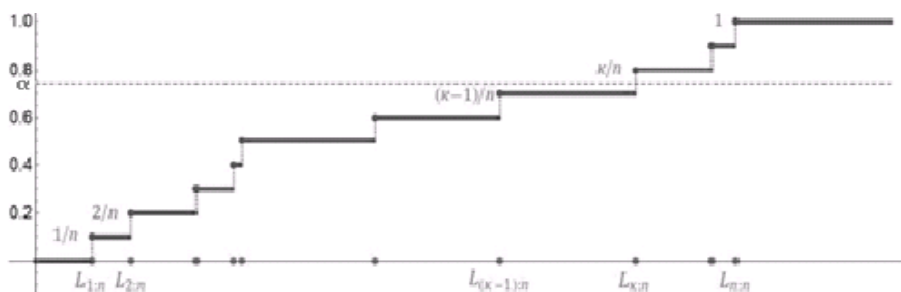


(Το σχήμα προέρχεται από το σύγγραμμα “Εισαγωγή στην διαχείριση Κινδύνων” Μπούτσικας Μιχαήλ (2020))

Η εμπειρική συνάρτηση κατανομής $\hat{F}_n(x)$ είναι αύξουσα, δεξιά συνεχής και έχει άλματα στα σημεία $L_{1:n}, L_{2:n}, \dots, L_{n:n}$, όπου $L_{1:n} \leq L_{2:n} \leq \dots \leq L_{n:n}$ είναι οι διατεταγμένες παρατηρήσεις $L_1, L_2, L_3, \dots, L_n$.

Ισχύει

$$\hat{F}_n(L_{\kappa:n}) = \frac{1}{n} \sum_{i=1}^n I_{[L_i \leq L_{\kappa}]} = \frac{\kappa}{n}, \quad \kappa = 1, 2, \dots, n$$



(Το σχήμα προέρχεται από το σύγγραμμα “Εισαγωγή στην διαχείριση Κινδύνων” Μπούτσικας Μιχαήλ (2020))

Με βάση τον ορισμό του $\text{VaR} = \inf \{x \in \mathbb{R} : F_L(x) \geq \alpha\}$ και αντικαθιστώντας το $F_L(x)$ με $\hat{F}_n(x)$, μπορούμε να εκτιμήσουμε το μέτρο κινδύνου.

$$\widehat{\text{VaR}}_{\alpha}(L) = \inf \{x \in \mathbb{R} : \hat{F}_n(x) \geq \alpha\} = L_{\kappa:n}, \quad \frac{\kappa - 1}{n} < \alpha \leq \frac{\kappa}{n}$$

Ισοδύναμα,

$$\widehat{\text{VaR}}_{\alpha}(L) = L_{\lceil n\alpha \rceil : n}$$

όπου $[x]$ είναι ο μικρότερος ακέραιος που είναι $\geq x$.

Με βάση τον ορισμό του Expected Shortfall της ζημιάς L σε επίπεδο εμπιστοσύνης $\alpha \in (0,1)$, $ES_\alpha(L) = E(L | L \geq VaR_\alpha(L))$, το ES μπορεί να εκτιμηθεί από τον δειγματικό μέσο των παρατηρήσεων $L_1, L_2, L_3, \dots, L_n$ (ισχύει $L_i \geq \widehat{VaR}_\alpha(L) = L_{[n\alpha]:n}$),

$$\widehat{ES}_\alpha(L) = \frac{L_{[n\alpha]:n} + \dots + L_{n:n}}{n - [n\alpha] + 1}$$

Η μέθοδος της εμπειρικής εκτίμησης από ιστορικά δεδομένα είναι απλή στην εφαρμογή της, ενώ δεν χρειάζεται να εφαρμοστούν υποθέσεις σχετικά με την εξάρτηση και την κατανομή των παρατηρήσεων. Ωστόσο, η αποτελεσματικότητα αυτής της μεθόδου εξαρτάται αρκετά από την διαθεσιμότητα και την επάρκεια των δεδομένων. Μεταβολές και ελλείψεις στα ιστορικά δεδομένα οδηγούν σε ανακριβείς εκτιμήσεις των μέτρων Var και ES. Επιπλέον, η διαθεσιμότητα ενός μεγάλου δείγματος είναι αναγκαία για την εξασφάλιση εύρους ακραίων ιστορικών σεναρίων τα οποία συμβάλουν σε καλύτερες εκτιμήσεις της ουρά της κατανομής ζημιών. Επομένως, η διαθεσιμότητα ενός μεγάλου δείγματος ιστορικών δεδομένων είναι απαραίτητη για την λήψη αξιόπιστων εκτιμήσεων των μέτρων κινδύνου VaR και ES.

1.2.3 Μέθοδος Variance-Covariance

Η ζημία που αντιστοιχεί την χρονική περίοδο i σε ένα χαρτοφυλάκιο, π.χ. έστω μετοχών, με σύνθεση $\underline{w} = (w_1, w_2, \dots, w_k)'$ (θεωρούμε ότι η σύνθεση παραμένει σταθερή σε αυτό το χρονικό διάστημα), δίνεται από την σχέση

$$\begin{aligned} L_i &= -(V_i - V_{i-1}) = -\left(\sum_{\ell=1}^k w_\ell S_i^{(\ell)} - \sum_{\ell=1}^k w_\ell S_{i-1}^{(\ell)}\right) = -\left(\sum_{\ell=1}^k w_\ell (S_i^{(\ell)} - S_{i-1}^{(\ell)})\right) \\ &= -\left(\sum_{\ell=1}^k w_\ell S_{i-1}^{(\ell)} R_i^{(\ell)}\right) \end{aligned}$$

$i = 1, 2, \dots$ όπου,

V_i : η αξία του χαρτοφυλακίου την χρονική περίοδο i

$S_i^{(\ell)}$: η αξία της μετοχής ℓ την χρονική περίοδο i

w_ℓ το πλήθος μετοχών ℓ στο χαρτοφυλάκιο

$R_i^{(\ell)}$: η ποσοστιαία μεταβολή (απόδοση) της αξίας της μετοχής ℓ την χρονική περίοδο i ,

$$R_i^{(\ell)} = \frac{S_i^{(\ell)} - S_{i-1}^{(\ell)}}{S_{i-1}^{(\ell)}} = e^{\ln(S_i^{(\ell)}/S_{i-1}^{(\ell)})} - 1 = e^{W_i^{(\ell)}} - 1$$

όπου $W_i^{(\ell)} = \ln(S_i^{(\ell)}/S_{i-1}^{(\ell)})$. Σύμφωνα με το μοντέλο Black and Scholes υποθέτουμε ότι τα τυχαία διανύσματα W_i ακολουθούν πολυδιάστατη κανονική κατανομή

$$W_i = (W_i^{(1)}, W_i^{(2)}, \dots, W_i^{(k)})' \sim N(\underline{\mu}, \Sigma)$$

και είναι μεταξύ τους ανεξάρτητα. Επομένως, οι αποδόσεις $R_i^{(1)}, R_i^{(2)}, \dots, R_i^{(k)}$ των k ζημιών ακολουθούν λογαριθμοκανονική κατανομή (μετατοπισμένη κατά 1, αφού $R_i^{(\ell)} = e^{W_i^{(\ell)}} - 1$) και είναι μεταξύ τους εξαρτημένες.

Στην περίπτωση που η χρονική περίοδος που εξετάζουμε είναι πολύ μικρή, για παράδειγμα ισούται με h και εκφράζει μία χρηματιστηριακή ημέρα του έτους, τότε το $W_i^{(\ell)}$ θα είναι κοντά στο 0. Επομένως, εάν βασιστούμε στην σχέση $e^x \approx x + 1$ που ισχύει για x μικρό, θα έχουμε προσεγγιστικά

$$R_i^{(\ell)} = \frac{S_i^{(\ell)} - S_{i-1}^{(\ell)}}{S_{i-1}^{(\ell)}} = e^{W_i^{(\ell)}} - 1 \approx W_i^{(\ell)}$$

Συνεπώς, μπορούμε να θεωρήσουμε για την i χρονική περίοδο για το διάνυσμα των αποδόσεων των μετοχών ότι

$$R_i = (R_i^{(1)}, R_i^{(2)}, \dots, R_i^{(k)})' \approx_d (W_i^{(1)}, W_i^{(2)}, \dots, W_i^{(k)})' = W_i \sim N_k(\underline{\mu}, \Sigma)$$

Οπότε, τα διανύσματα $R_i, i = 1, 2, \dots$ είναι ανεξάρτητα μεταξύ τους.

Σε γενικότερες εφαρμογές όπου η τ.μ. $S_i^{(\ell)}$ εκφράζει την αξία της ℓ επένδυσης (επί 1 χρηματικής μονάδας) την i χρονική περίοδο και w_ℓ είναι το ποσό που έχει επενδυθεί στην ℓ επένδυση, μπορούμε απευθείας να θεωρήσουμε ότι το διάνυσμα των αποδόσεων $R_i \sim N_k(\underline{\mu}, \Sigma)$.

Σύμφωνα με τα παραπάνω καταλήγουμε ότι η ζημία του χαρτοφυλακίου για την i χρονική περίοδο ισούται με τον γραμμικό συνδυασμό τ.μ. (τυχαίων μεταβλητών) που ακολουθούν πολυδιάστατη κανονική κατανομή,

$$L_i = - \left(\sum_{\ell=1}^k w_{\ell} S_{i-1}^{(\ell)} R_i^{(\ell)} \right) = - \left(\sum_{\ell=1}^k u_i^{(\ell)} R_i^{(\ell)} \right) = - \underline{u}'_i * R_i,$$

όπου θέτουμε $u_i^{(\ell)} = w_{\ell} S_{i-1}^{(\ell)}$, $\underline{u}'_i = (u_i^{(1)}, \dots, u_i^{(k)})'$.

Συνεπώς, η L_i ακολουθεί κανονική κατανομή με μέση τιμή

$$E(L_i) = E \left(- \sum_{\ell=1}^k u_i^{(\ell)} R_i^{(\ell)} \right) = - \sum_{\ell=1}^k u_i^{(\ell)} \mu_{\ell} = - \underline{u}'_i * \underline{\mu}$$

και διασπορά

$$\begin{aligned} V(L_i) &= V \left(- \sum_{\ell=1}^k u_i^{(\ell)} R_i^{(\ell)} \right) = V \left(\sum_{\ell=1}^k u_i^{(\ell)} R_i^{(\ell)} \right) = \text{Cov} \left(\sum_{\ell=1}^k u_i^{(\ell)} R_i^{(\ell)}, \sum_{m=1}^k u_i^{(m)} R_i^{(m)} \right) \\ &= \sum_{\ell=1}^k \sum_{m=1}^k u_i^{(\ell)} u_i^{(m)} \text{Cov}(R_i^{(\ell)}, R_i^{(m)}) = \sum_{\ell=1}^k \sum_{m=1}^k u_i^{(\ell)} u_i^{(m)} \sigma_{\ell m} = \underline{u}'_i \Sigma \underline{u}_i \end{aligned}$$

Γιατί για $k = 2$ ισχύει

$$\begin{aligned} \underline{u}'_i \Sigma \underline{u}_i &= \begin{pmatrix} u_i^{(1)} & u_i^{(2)} \end{pmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \begin{pmatrix} u_i^{(1)} \\ u_i^{(2)} \end{pmatrix} \\ &= u_i^{(1)} u_i^{(1)} \sigma_{11} + u_i^{(1)} u_i^{(2)} \sigma_{12} + u_i^{(2)} u_i^{(1)} \sigma_{21} + u_i^{(2)} u_i^{(2)} \sigma_{22} \\ &= \sum_{\ell=1}^2 \sum_{m=1}^2 u_i^{(\ell)} u_i^{(m)} \sigma_{\ell m} \end{aligned}$$

Καταλήγουμε ότι η ζημία της i περιόδου L_i ακολουθεί κανονική κατανομή

$$L_i = - \underline{u}'_i * R_i \sim N(- \underline{u}'_i \underline{\mu}, \underline{u}'_i \Sigma \underline{u}_i)$$

Καθώς, έχουμε δει ότι αν $L \sim N(\mu, \sigma^2)$ τότε $\text{VaR}_\alpha(L) = \mu + \sigma\Phi^{-1}(\alpha)$ και $\text{ES}_\alpha(L) = \mu + \sigma \frac{\varphi(\Phi^{-1}(\alpha))}{1-\alpha}$. Συνεπώς, ο κίνδυνος του χαρτοφυλακίου την χρονική περίοδο i (δεδομένου ότι βρισκόμαστε στην αρχή της περιόδου και γνωρίζουμε τις τιμές $S_{i-1}^{(\ell)}$) ισούται με

$$\text{VaR}_\alpha(L_i) = -u'_i \mu + \sqrt{u'_i \Sigma u_i} \Phi^{-1}(\alpha)$$

$$\text{ES}_\alpha(L_i) = -u'_i \mu + \sqrt{u'_i \Sigma u_i} \frac{\varphi(\Phi^{-1}(\alpha))}{1-\alpha}$$

όπου $u_i^{(\ell)} = w_\ell S_{i-1}^{(\ell)}$, $\ell = 1, 2, \dots, k$

Η μέθοδος Variance-Covariance προσφέρει μια απλή και αναλυτική λύση εκτίμησης των μέτρων κινδύνου Var και ES. Όμως, η υπόθεση της κανονικότητας δεν είναι πάντα ρεαλιστική και ικανοποιητική προσέγγιση οδηγώντας συνήθως σε υποεκτίμηση του κινδύνου. Μία καλύτερη προσέγγιση μπορεί να επιτευχθεί υποθέτοντας ελλειπτικές πολυδιάστατες κατανομές οι οποίες ενδέχεται να ταιριάζουν καλύτερα στα δεδομένα.

Για περισσότερες πληροφορίες για την εκτίμηση των Μέτρων κινδύνου ο ενδιαφερόμενος αναγνώστης μπορεί να ανατρέξει στα συγγράμματα. Denuit, J. Dhaene, M. Goovaerts, R. Kaas (2005), Alexander J. McNeil, Rudiger Frey, Paul Embrechts (2005) και M. Μπούτσικας (2020).

Κεφάλαιο 2

Συναρτήσεις Σύζευξης

Ένα βασικό βήμα στην ανάλυση κινδύνου είναι η δημιουργία ενός στατιστικού μοντέλου το οποίο απεικονίζει την τυχαιότητα που είναι εγγενής σε μία κατάσταση. Δεδομένης της πολυπλοκότητας των μοντέλων αυτών, τα στοχαστικά αποτελέσματα μετρώνται σε πολλές διαστάσεις και η κατανόηση των σχέσεων και της εξάρτησης μεταξύ των μονοδιάστατων αποτελεσμάτων ενός πολυδιάστατου μοντέλου αποτελεί βασικό προβληματισμό στην Στατιστική επιστήμη.

Για μεγάλο χρονικό διάστημα η στατιστική μοντελοποίηση στην αναλογιστική επιστήμη και στα χρηματοοικονομικά βασιζόταν κυρίως σε απλουστευμένες παραδοχές. Συγκεκριμένα, η κανονική κατανομή κυριάρχησε στην μελέτη των πολυδιάστατων κατανομών κυρίως λόγω της εύκολης μαθηματικής επεξεργασίας και δυνατότητας διαχείρισης. Ωστόσο, σε κείμενα σχετικά με την μελέτη πολυδιάστατων κατανομών, (π.χ. βλ. Krzanowski (1988)), είχε αρχίσει να αναγνωρίζεται η ανάγκη εξέτασης εναλλακτικών προσεγγίσεων καθώς η χρήση της κανονικής κατανομής σε αναλογιστικές εφαρμογές δεν παρέχει επαρκή προσέγγιση των δεδομένων.

Στο κεφάλαιο αυτό θα ασχοληθούμε με την μοντελοποίηση της εξάρτησης πολυδιάστατων αποτελεσμάτων χρησιμοποιώντας την έννοια των συναρτήσεων σύζευξης (Copulas), ένα εργαλείο για την κατανόηση των σχέσεων μεταξύ τυχαίων μεταβλητών. Η μαθηματική έκφραση των συναρτήσεων σύζευξης παρουσιάστηκε για πρώτη φορά το 1959 από τον Sklar μέσω του θεωρήματος του το οποίο φέρει το όνομα του, ενώ η διαθέσιμη βιβλιογραφία σχετικά με τις στατιστικές ιδιότητες και τις εφαρμογές των συζεύξεων αναπτύσσεται ραγδαία τα τελευταία χρόνια.

Η συνάρτηση σύζευξης (Copula) συνδέει μονοδιάστατες περιθώριες συναρτήσεις κατανομής με την από κοινού πολυδιάστατη κατανομή τους. Συγκεκριμένα, επιλέγοντας περιθώριες κατανομές οι οποίες ταιριάζουν στα διαθέσιμα δεδομένα, είναι εφικτό να προσδιοριστεί η από κοινού πολυδιάστατη κατανομή τους μέσω μιας κατάλληλης συνάρτησης σύζευξης. Επιπλέον, όπως προκύπτει από το βασικό θεώρημα του Sklar, ο ορισμός της συνάρτησης σύζευξης δεν θέτει περιορισμούς στην επιλογή των περιθωρίων κατανομών.

Στο κεφάλαιο αυτό θα περιγράψουμε βασικές ιδιότητες των Copulas και θα παρουσιάσουμε οικογένειες των συναρτήσεων αυτών οι οποίες έχουν εμφανιστεί κατά καιρούς στην βιβλιογραφία. Σε όλη την έκταση του κεφαλαίου θεωρούμε ότι οι περιθώριες κατανομές είναι συνεχής. Τα α-

ποτελέσματα από την μοντελοποίηση της εξάρτησης μεταξύ διακριτών κατανομών δεν είναι εξίσου αποτελεσματικά, ενώ μπορεί να υπάρχουν περισσότερες από μία copula συνάρτηση για τον προσδιορισμό της από κοινού πολυδιάστατης συνάρτησης κατανομής. Επίσης, για λόγους απλότητας θα εστιάσουμε κυρίως σε διδιάστατες κατανομές ώστε να αποφύγουμε την πολυπλοκότητα φορμαλισμού των πολυδιάστατων κατανομών.

Για περισσότερες πληροφορίες και λεπτομέρειες σχετικά με τις συναρτήσεις σύζευξης - Copula ο αναγνώστης μπορεί να συμβουλευτεί τα συγγράμματα M. Denuit, J. Dhaene, M. Goovaerts, R. Kaas (2005) και Edward W. Frees, Emiliano A. Valdez (1998).

2.1 Συναρτήσεις Σύζευξης (Copulas)

Δοθείσης μιας διδιάστατης συνάρτησης κατανομής F_X με περιθώριες συναρτήσεις F_1 και F_2 , μπορούμε για κάθε ζεύγος πραγματικών αριθμών $x = (x_1, x_2)$ να συνδέσουμε τα τρία αυτά μεγέθη. Συγκεκριμένα, κάθε ζεύγος (x_1, x_2) πραγματικών αριθμών οδηγεί σε ένα σημείο

$$(F_1(x_1), F_2(x_2))$$

το οποίο με την σειρά του αντιστοιχεί σε έναν αριθμό $F_X(x_1, x_2)$ στο διάστημα $[0,1]$. Θα δούμε στην συνέχεια ότι η απεικόνιση η οποία συνδέει την τιμή της από κοινού συνάρτησης κατανομής με ένα διατεταγμένο ζεύγος περιθωρίων κατανομών αποτελεί μια συνάρτηση και συγκεκριμένα την συνάρτηση σύζευξης - Copula.

Ορισμός: Η διδιάστατη copula C είναι μια απεικόνιση από τον χώρο $[0,1]^2 = [0,1] \times [0,1]$ στο διάστημα $[0,1]$, η οποία είναι μη φθίνουσα, δεξιά συνεχής και ικανοποιεί τις παρακάτω υποθέσεις:

- $\lim_{u_i \rightarrow 0} C(u_1, u_2) = 0$ για $i = 1, 2$
- $\lim_{u_1 \rightarrow 1} C(u_1, u_2) = u_2$ και $\lim_{u_2 \rightarrow 1} C(u_1, u_2) = u_1$
- Η ανισότητα

$$C(v_1, v_2) - C(u_1, v_2) - C(v_1, u_2) + C(u_1, u_2) \geq 0$$

ισχύει για $u_1 \leq v_1, u_2 \leq v_2$

Ουσιαστικά η συνάρτηση C θα πρέπει να έχει τις ιδιότητες μιας διδιάστατης συνάρτησης κατανομής στο σύνολο $[0,1] \times [0,1]$ με ομοιόμορφες στο $[0,1]$ περιθώριες.

2.1.1 Το Θεώρημα Sklar

Το θεώρημα του Sklar αποτελεί βασικό θεώρημα στην θεωρία των συζεύξεων και θεμέλιο σε πολλές εφαρμογές, καθώς διευκρινίζει τον ρόλο των συναρτήσεων σύζευξης στην σχέση που υπάρχει μεταξύ των πολυδιάστατων σ.κ. και των μονοδιάστατων περιθωρίων συναρτήσεων τους.

Θεώρημα Sklar. Έστω συνάρτηση F_X με συνεχείς περιθώριες συναρτήσεις κατανομών F_1 και F_2 . Τότε υπάρχει μοναδική συνάρτηση σύζευξης (copula) C τέτοια ώστε για κάθε $x = (x_1, x_2) \in \mathbb{R}^2$ ισχύει,

$$F_X(x_1, x_2) = C(F_1(x_1), F_2(x_2)) \quad (1)$$

Αντιστρόφως, εάν C είναι μια συνάρτηση σύζευξης και F_1, F_2 συναρτήσεις κατανομής τότε η συνάρτηση F_X , όπως ορίζεται από την σχέση (1), είναι διδιάστατη συνάρτηση κατανομής με περιθώριες συναρτήσεις F_1 και F_2 .

Η σχέση (1) συνδέει τις περιθώριες συναρτήσεις F_1, F_2 με την από κοινού σ.κ. F_X . Η δομή εξάρτησης περιγράφεται από την συνάρτηση C , συνεπώς ο τρόπος με τον οποίο αλληλεπιδρούν οι τυχαίες μεταβλητές X_1, X_2 αποδίδεται από την copula ανεξαρτήτως της κλίμακας στην οποία μετρώνται οι τυχαίες μεταβλητές.

Μία ιδιότητα η οποία χρησιμοποιείται αρκετά στην μελέτη των συναρτήσεων σύζευξης είναι αυτή του μετασχηματισμού πιθανότητας. Συγκεκριμένα, αν η τ.μ. X έχει συνεχή συνάρτηση κατανομής F_X , τότε ισχύει $F_X(X) \sim U(0,1)$.

Πράγματι, για κάθε $0 < u < 1$,

$$P[F_X(X) \geq u] = P[X \geq F_X^{-1}(u)] = 1 - F_X(F_X^{-1}(u)) = 1 - u$$

άρα, $F_X(X) \sim U(0,1)$.

Επιπλέον, με βάση την παραπάνω ιδιότητα αποδεικνύεται ότι, εάν F_1, F_2 είναι περιθώριες συνεχείς συναρτήσεις τότε

$$C(u_1, u_2) = F_X(F_1^{-1}(u_1), F_2^{-1}(u_2)), \quad (u_1, u_2) \in [0, 1]^2$$

Παρατηρούμε ότι, η copula εκφράζει την εξάρτηση σε ποσοστιαία κλίμακα καθώς ισούται με την από κοινού σ.κ. του διανύσματος X .

Στην περίπτωση που οι περιθώριες συναρτήσεις κατανομής F_1 και F_2 δεν είναι συνεχείς τότε το θεώρημα Sklar διασφαλίζει την ύπαρξη της συνάρτησης σύζευξης του F_X , ωστόσο, δεν εξασφαλίζει την μοναδικότητα της συνάρτησης αυτής.

Παρακάτω παρουσιάζουμε μερικά βασικά παραδείγματα και ειδικές περιπτώσεις της σχέσης (1). Συγκεκριμένα,

Η Ανεξάρτητη σύζευξη C_I (Independed Copula C_I). Εάν οι τ.μ. X_1 και X_2 είναι ανεξάρτητες τυχαίες μεταβλητές με αντίστοιχες συναρτήσεις κατανομής F_1 και F_2 , τότε η από κοινού συνάρτηση κατανομής ισούται με $F_X(x) = F_1(x_1)F_2(x_2)$ και η συνάρτηση copula δίνεται από την σχέση

$$C_I(u_1, u_2) = u_1 \cdot u_2, \quad (u_1, u_2) \in [0, 1]^2$$

εάν οι μεταβλητές X_1 και X_2 ικανοποιούν την σχέση (1) (Θεώρημα Sklar), τότε είναι ανεξάρτητες αν και μόνο αν $C \equiv C_I$.

Άνω όριο Fréchet Copula C_U (The Fréchet upper bound Copula C_U). Η συνάρτηση σύζευξης που σχετίζεται με το άνω όριο Fréchet συμβολίζεται με C_U και ισούται με

$$C_U(u_1, u_2) = \min\{u_1, u_2\}, \quad (u_1, u_2) \in [0, 1]^2$$

Κάτω όριο Fréchet Copula C_L (The Fréchet lower bound Copula C_L). Το κάτω όριο Fréchet της Copula συμβολίζεται με C_L και δίνεται από την σχέση

$$C_L(u_1, u_2) = \max \{0, u_1 + u_2 - 1\}, \quad (u_1, u_2) \in [0, 1]^2$$

Με βάση τα προαναφερθέντα όρια Fréchet, για κάθε συνάρτηση σύζευξης αποδεικνύεται ότι ισχύει η ανισότητα

$$C_L(u_1, u_2) \leq C(u_1, u_2) \leq C_U(u_1, u_2), \quad (u_1, u_2) \in [0, 1]^2$$

Επομένως, κάθε συνάρτηση σύζευξης παρεμβάλλεται μεταξύ του άνω και του κάτω ορίου Fréchet της Copula.

Επιπλέον, μια συνάρτηση σύζευξης λέγεται πλήρης (Comprehensive Copula) όταν ικανοποιεί τις ειδικές περιπτώσεις της ανεξάρτητης copula και του άνω και κάτω ορίου Fréchet (δηλαδή για συγκεκριμένες τιμές των παραμέτρων της συμπίπτει με τις C_U, C_L, C_I)

2.1.2 Σύνδεση Copula Συναρτήσεων και Συναρτήσεων Πυκνότητας Πιθανότητας

Κάτω από κατάλληλες συνθήκες, η από κοινού συνάρτηση πυκνότητας πιθανότητας του διδιάστατου διανύσματος $X = (X_1, X_2)$ μπορεί να γραφτεί ως γινόμενο των περιθωρίων συναρτήσεων πυκνότητας πιθανότητας και της συνάρτησης πυκνότητας σύζευξης (copula d.f.). Η συνάρτηση πυκνότητας σύζευξης εμπεριέχει και αποτυπώνει όλη την πληροφορία σχετικά με την εξάρτηση που υπάρχει μεταξύ των τυχαίων μεταβλητών X_1, X_2 .

Συγκεκριμένα, εάν οι περιθώριες συναρτήσεις κατανομής F_1 και F_2 είναι συνεχείς με αντίστοιχες συναρτήσεις πυκνότητας πιθανότητας f_1 και f_2 , τότε η από κοινού σ.π.π. του X μπορεί να γραφτεί ως εξής

$$\begin{aligned} f_X(x_1, x_2) &= \frac{\partial^2}{\partial u_1 \partial u_2} F_X(x_1, x_2) = \frac{\partial^2}{\partial u_1 \partial u_2} C(F_1(x_1), F_2(x_2)) \\ &= f_1(x_1) f_2(x_2) c(F_1(x_1), F_2(x_2)), \quad x \in \mathbb{R}^2 \end{aligned} \quad (2)$$

όπου η συνάρτηση πυκνότητας σύζευξης c δίνεται από την σχέση

$$c(u_1, u_2) = \frac{\partial^2}{\partial u_1 \partial u_2} C(u_1, u_2), \quad (u_1, u_2) \in [0, 1]^2$$

Στην σχέση (2), ο όρος $f_1(x_1)f_2(x_2)$ ισούται με την από κοινού σ.π.π. του τυχαίου διανύσματος X σε περίπτωση ανεξαρτησίας των X_i , ενώ η δεύτερη μερική παράγωγος της συνάρτησης C μπορεί να ερμηνευτεί ως μέτρο εξάρτησης των X_i (όταν υπάρχει). Επομένως, η από κοινού σ.π.π. στο σημείο $x = (x_1, x_2)$ ισούται με την σ.π.π. που αντιστοιχεί στην περίπτωση ανεξαρτησίας των X_i πολλαπλασιασμένη με τον όρο $c(F_1(x_1), F_2(x_2))$. Μέσω του τελευταίου παράγοντα, αλλοιώνεται η ανεξαρτησία και εισέρχεται η πραγματική δομή εξάρτησης, επομένως, η από κοινού σ.π.π. στο σημείο x προκύπτει από την ανεξάρτητη σ.π.π. $f_1(x_1)f_2(x_2)$ σταθμισμένη κατά τον όρο $c(F_1(x_1), F_2(x_2))$.

Όσον αφορά την ύπαρξη των μερικών παραγώγων της συνάρτησης C , αποδεικνύεται η εξής πρόταση:

Πρόταση. Έστω C μια συνάρτηση σύζευξης copula. Για κάθε $u_2 \in [0, 1]$ η μερική παράγωγος $\frac{\partial}{\partial u_1} C(u_1, u_2)$ υπάρχει σχεδόν παντού και για κάθε ζεύγος (u_1, u_2) έχουμε

$$0 \leq \frac{\partial}{\partial u_1} C(u_1, u_2) \leq 1$$

Αντίστοιχα, για κάθε $u_1 \in [0, 1]$ η μερική παράγωγος $\frac{\partial}{\partial u_2} C(u_1, u_2)$ υπάρχει σχεδόν παντού και για κάθε ζεύγος (u_1, u_2) έχουμε

$$0 \leq \frac{\partial}{\partial u_2} C(u_1, u_2) \leq 1$$

όπου σχεδόν παντού σημαίνει οπουδήποτε εκτός από ένα σύνολο μέτρου 0.

Επιπλέον, η μερική παράγωγος της συνάρτησης σύζευξης συνδέεται με την δεσμευμένη συνάρτηση ως εξής:

$$C_{2|1}(u_2|u_1) = P(U_2 \leq u_2 | U_1 = u_1) = \lim_{\delta \rightarrow 0} \frac{C(u_1 + \delta, u_2) - C(u_1, u_2)}{\delta} = \frac{\partial}{\partial u_1} C(u_1, u_2)$$

και όμοια

$$C_{1|2}(u_1|u_2) = \frac{\partial}{\partial u_2} C(u_1, u_2)$$

Με βάση την σχέση (1) (θεώρημα Sklar), για κάθε $x = (x_1, x_2) \in \mathbb{R}^2$ ισχύει

$$\begin{aligned} P[X_2 \leq x_2 | X_1 = x_1] &= \frac{\frac{\partial}{\partial x_1} F_X(x_1, x_2)}{f_{X_1}(x_1)} = \frac{\frac{\partial}{\partial x_1} C(F_1(x_1), F_2(x_2))}{f_{X_1}(x_1)} \\ &= \frac{f_{X_1}(x_1) \frac{\partial}{\partial u_1} C(F_1(x_1), F_2(x_2))}{f_{X_1}(x_1)} = C_{2|1}(F_2(x_2) | F_1(x_1)) \end{aligned}$$

και όμοια

$$P[X_1 \leq x_1 | X_2 = x_2] = C_{1|2}(F_1(x_1) | F_2(x_2))$$

2.2 Οικογένειες Διδιάστατων συναρτήσεων σύζευξης

Στην παρακάτω παράγραφο θα δούμε μερικές από τις πιο γνωστές διδιάστατες συναρτήσεις σύζευξης.

Clayton Copula (σύζευξη Clayton)

Η Clayton σύζευξη αποδίδεται συνήθως στον Clayton (1978) αλλά η προέλευση της μπορεί στην πραγματικότητα να εντοπιστεί στους Kimeldorf και Sampson (1975). Η συνάρτηση δίνεται από τον τύπο

$$C_a(u_1, u_2) = (u_1^{-a} + u_2^{-a} - 1)^{-1/a}, \quad a > 0$$

όπου $(u_1, u_2) \in [0, 1]^2$

Σύμφωνα με μελέτες (Cook και Johnson (1981)) η παράμετρος a ερμηνεύεται ως μέτρο έντασης της εξάρτησης μεταξύ των u_i .

Για τις ακραίες περιπτώσεις ισχύει ότι

$$\lim_{a \rightarrow +\infty} C_a(u_1, u_2) = \min\{u_1, u_2\} = C_U(u)$$

η σχέση εξάρτησης είναι η μέγιστη όταν το a τείνει στο άπειρο, ενώ

$$\lim_{a \rightarrow 0} C_a(u_1, u_2) = u_1 u_2 = C_I(u)$$

η ανεξαρτησία μεταξύ των u_i επιτυγχάνεται όταν το a τείνει στο 0.

Οι δεσμευμένες συναρτήσεις προκύπτουν από τις μερικές παραγώγους της Clayton copula. Επομένως, οι $C_{1|2}$ και $C_{2|1}$ είναι ίσες με

$$C_{2|1}(u_2|u_1) = \frac{\partial}{\partial u_1} C_a(u) = (1 + u_1^a(u_2^{-a} - 1))^{-1 - \frac{1}{a}}$$

αντίστοιχα προκύπτει για $C_{1|2}$.

Επιπλέον, η συνάρτηση πυκνότητας πιθανότητας που σχετίζεται με την Clayton Copula δίνεται από την σχέση

$$c_a(u_1, u_2) = \frac{1+a}{(u_1 u_2)^{a+1}} (u_1^{-a} + u_2^{-a} - 1)^{-2 - \frac{1}{a}}$$

Frank Copula (σύζευξη Frank)

Η συγκεκριμένη οικογένεια Copula μελετήθηκε από τον Genest (1987), ενώ εμφανίστηκε για πρώτη φορά στο άρθρο του Frank (1979) ως λύση σε μία εξίσωση. Η συνάρτηση δίνεται από τον τύπο

$$C_a(u_1, u_2) = -\frac{1}{a} \ln \left(1 + \frac{(\exp(-au_1) - 1)(\exp(-au_2) - 1)}{\exp(-a) - 1} \right), \quad a \neq 0$$

όπου $(u_1, u_2) \in [0, 1]^2$.

Άμεσα αποδεικνύονται οι παρακάτω ειδικές περιπτώσεις για την συνάρτηση C_a ,

$$\lim_{a \rightarrow -\infty} C_a = C_L, \quad \lim_{a \rightarrow +\infty} C_a = C_U, \quad \lim_{a \rightarrow 0} C_a = C_I$$

Επομένως, η Frank Copula είναι πλήρης (Comprehensive Copula) καθώς ικανοποιεί την ανεξάρτητη copula και το άνω και το κάτω όριο Fréchet.

Οι δεσμευμένες συναρτήσεις δίνονται από την σχέση

$$C_{2|1}(u_2|u_1) = \frac{\exp(-au_2) - \exp(-a(u_1 + u_2))}{1 - \exp(-a) - (1 - \exp(-au_1))(1 - \exp(-au_2))}$$

όπου αντίστοιχα προκύπτει για $C_{1|2}$.

Επιπλέον, η συνάρτηση πυκνότητας πιθανότητας που σχετίζεται με την Frank Copula ισούται με

$$c_a(u_1, u_2) = \frac{a \exp(-a(u_1 + u_2))(1 - \exp(-a))}{(\exp(-a(u_1 + u_2)) - \exp(-au_1) - \exp(-au_2) + \exp(-a))^2}$$

Gumbel Copula (σύζευξη Gumbel)

Η συνάρτηση σύζευξης Gumbel δίνεται από τον παρακάτω κλειστό τύπο

$$C_\alpha(u_1, u_2) = \exp\{-((- \ln u_1)^\alpha + (- \ln u_2)^\alpha)^{1/\alpha}\}, \quad 1 \leq \alpha < \infty$$

όπου η παράμετρος α αναπαριστά την ένταση της εξάρτησης. Εάν, $\alpha = 1$ τότε καταλήγουμε στην ειδική περίπτωση της ανεξάρτητης Copula, ενώ όταν $\alpha \rightarrow \infty$ προκύπτει το άνω όριο Fréchet της σύζευξης. Επομένως, η Gumbel Copula παρεμβάλλεται μεταξύ της ανεξάρτητης Copula και του άνω ορίου Fréchet.

Οι Clayton, Frank και Gumbel copulas ανήκουν στην οικογένεια της Αρχιμήδειας Copula την οποία θα εξετάσουμε προς το τέλος του συγκεκριμένου κεφαλαίου. Σε αντίθεση με τις τρεις προαναφερθείσες συναρτήσεις, υπάρχουν οικογένειες των οποίων οι συναρτήσεις σύζευξης δεν αποτελούν κλειστό τύπο αλλά προκύπτουν μέσω γνωστών σ.κ..

Κανονική Copula (Normal Copula ή Gaussian Copula)

Η κανονική copula ή Gaussian copula αποτυπώνει την εξάρτηση που περιέχει ένα διδιάστατο τυχαίο διάνυσμα το οποίο ακολουθεί κανονική κατανομή. Η συνάρτηση C δίνεται από την σχέση

$$C_\rho(u_1, u_2) = H_\rho(\Phi^{-1}(u_1), \Phi^{-1}(u_2)), \quad \rho \in (-1, 1)$$

όπου Φ η συνάρτηση κατανομής της μονοδιάστατης τυποποιημένης $N(0,1)$ κατανομής και H_ρ η σ.κ. της διδιάστατης κανονικής κατανομής με συντελεστή συσχέτισης ρ . Πιο αναλυτικά,

$$C_\rho(u_1, u_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \exp\left(\frac{-(\xi_1^2 - 2\rho\xi_1\xi_2 + \xi_2^2)}{2(1-\rho^2)}\right) d\xi_1 d\xi_2$$

Για τις ειδικές περιπτώσεις για την συνάρτηση C_ρ ισχύει,

$$\lim_{\rho \rightarrow -1} C_\rho = C_L, \quad \lim_{\rho \rightarrow 0} C_\rho = C_U, \quad \lim_{\rho \rightarrow 1} C_\rho = C_I$$

Η αντίστοιχη συνάρτηση πυκνότητας πιθανότητας της συνάρτησης σύζευξης προκύπτει ως εξής

$$c_\rho(u_1, u_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(\frac{-(\zeta_1^2 - 2\rho\zeta_1\zeta_2 + \zeta_2^2)}{2(1-\rho^2)}\right) \frac{d}{du_1} \Phi^{-1}(u_1) \frac{d}{du_2} \Phi^{-1}(u_2)$$

όπου $\zeta_i = \Phi^{-1}(u_i)$ για $i = 1, 2$. Επιπλέον, η $\varphi = \Phi'$ αποτελεί την σ.π.π. της κανονικής κατανομής, άρα

$$\frac{d}{du_i} \Phi^{-1}(u_i) = \frac{1}{\varphi(\Phi^{-1}(u_i))} = \sqrt{2\pi} \exp\left(\frac{-\zeta_i^2}{2}\right)$$

Επομένως, καταλήγουμε στην σ.π.π. της κανονικής copula η οποία ισούται με

$$c_\rho(u_1, u_2) = \frac{1}{\sqrt{1-\rho^2}} \exp\left(\frac{-(\zeta_1^2 - 2\rho\zeta_1\zeta_2 + \zeta_2^2)}{2(1-\rho^2)}\right) \exp\left(\frac{\zeta_1^2 + \zeta_2^2}{2}\right)$$

και ισχύει ότι η ένταση της εξάρτησης μεταβάλλεται με το ρ . Δεδομένου ότι η κανονική copula μπορεί να γραφτεί ως εξής

$$C_\rho(u_1, u_2) = \int_0^{u_1} \Phi\left(\frac{\Phi^{-1}(u_2) - \rho\Phi^{-1}(t)}{\sqrt{1-\rho^2}}\right) dt$$

οι δεσμευμένες συναρτήσεις που σχετίζονται με την Gaussian copula είναι ίσες με:

$$C_{2|1}(u_2|u_1) = \Phi\left(\frac{\Phi^{-1}(u_2) - \rho\Phi^{-1}(u_1)}{\sqrt{1-\rho^2}}\right)$$

Student Copula

Όπως η κανονική copula σχετίζεται με την διδιάστατη κανονική κατανομή, αντίστοιχα και η student copula προκύπτει από μία διδιάστατη student κατανομή και ορίζεται ως εξής

$$\begin{aligned} C_{m,a}(u_1, u_2) &= t_{m,a}(t_m^{-1}(u_1), t_m^{-1}(u_2)) \\ &= \int_{-\infty}^{t_m^{-1}(u_1)} \int_{-\infty}^{t_m^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-a^2}} \left(1 + \frac{\xi_1^2 - 2a\xi_1\xi_2 + \xi_2^2}{m(1-a^2)}\right)^{-\frac{m+2}{2}} d\xi_1 d\xi_2 \end{aligned}$$

όπου m οι βαθμοί ελευθερίας της student κατανομής και $0 < a < 1$. Όταν οι βαθμοί ελευθερίας m τείνουν στο άπειρο $+\infty$ τότε η student copula συμπίπτει με την κανονική copula.

Να σημειώσουμε ότι η συνάρτηση κατανομής της μονοδιάστατης Student κατανομής με m βαθμούς ελευθερίας δίνεται από την σχέση

$$t_m(x) = \int_{-\infty}^x \frac{\Gamma\left(\frac{m+1}{2}\right)}{\sqrt{m\pi}\Gamma\left(\frac{m}{2}\right)} \left(1 + \frac{\xi^2}{m}\right)^{-\frac{m+1}{2}} d\xi$$

αντίστοιχα, η σ.κ. της διδιάστατης Student κατανομής ισούται με

$$t_{m,a}(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \frac{1}{2\pi\sqrt{1-a^2}} \left(1 + \frac{\xi_1^2 - 2a\xi_1\xi_2 + \xi_2^2}{m(1-a^2)}\right)^{-\frac{m+2}{2}} d\xi_1 d\xi_2$$

για $0 < a < 1$.

Η σ.π.π. της Student Copula δίνεται από την σχέση

$$c_{m,a}(u_1, u_2) = a^{-\frac{1}{2}} \frac{\Gamma\left(\frac{m+2}{2}\right) \Gamma\left(\frac{m}{2}\right)}{\left(\Gamma\left(\frac{m+1}{2}\right)\right)^2} \left(1 + \frac{\zeta_1^2 - 2a\zeta_1\zeta_2 + \zeta_2^2}{m(1-a^2)}\right)^{-\frac{m+2}{2}} \\ \times \left(1 + \frac{\zeta_1}{m}\right)^{-\frac{m+2}{2}} \left(1 + \frac{\zeta_2}{m}\right)^{-\frac{m+2}{2}}$$

όπου $\zeta_1 = t_m^{-1}(u_1)$ και $\zeta_2 = t_m^{-1}(u_2)$.

Επιπλέον, η δεσμευμένη Student copula δίνεται από την τύπο

$$C_{2|1}(u_2|u_1) = t_{m+1} \left(\sqrt{\frac{m+1}{m + (t_m^{-1}(u_1))^2} \frac{t_m^{-1}(u_2) - at_m^{-1}(u_1)}{\sqrt{1-a^2}}} \right)$$

Η οικογένεια Student copula ικανοποιεί το άνω και το κάτω όριο Fréchet για κάθε τιμή του m

$$\lim_{a \rightarrow -1} C_{m,a} = C_L, \quad \lim_{a \rightarrow 1} C_{m,a} = C_U$$

ωστόσο, για πεπερασμένη τιμή του m $C_{m,0} \neq C_I$.

Για περισσότερα αποτελέσματα σχετικά με τις διδιάστατες συναρτήσεις σύζευξης (Copulas) ο ενδιαφερόμενος αναγνώστης μπορεί να ανατρέξει στα συγγράμματα M. Denuit, J. Dhaene, M. Goovaerts, R. Kaas (2005) και Alexander J. McNeil, Rudiger Frey, Paul Embrechts (2005).

2.3 Ιδιότητες των συναρτήσεων σύζευξης (Copulas)

Η από κοινού συνάρτηση κατανομής ενός τυχαίου διανύσματος εμπεριέχει πληροφορία για τη συμπεριφορά των περιθωρίων κατανομών των τυχαίων μεταβλητών (κινδύνων) αλλά και για την μεταξύ τους εξάρτηση. Αντίθετα, η συνάρτηση σύζευξης επιτυγχάνει να απομονώνει και να αποτυπώνει μόνο την δομή εξάρτησης των τυχαίων μεταβλητών. Στο παρακάτω κεφάλαιο θα δούμε σημαντικές ιδιότητες και έννοιες που σχετίζονται με τις συναρτήσεις σύζευξης.

2.3.1 Σύζευξη επιβίωσης (Survival Copula)

Ορισμός: Εάν η C είναι μία Copula, τότε και η \bar{C} είναι copula και ορίζεται για $u \in [0, 1]^2$ ως εξής

$$\bar{C}(u_1, u_2) = C(1 - u_1, 1 - u_2) + u_1 + u_2 - 1$$

Η συνάρτηση \bar{C} καλείται copula επιβίωσης και σχετίζεται με την C . Προσοχή θα πρέπει να δοθεί ώστε να μην συγχέεται η copula επιβίωσης με την από κοινού συνάρτηση επιβίωσης δύο ομοιόμορφων $\text{Uni}(0,1)$ τυχαίων μεταβλητών (U_1, U_2) των οποίων η από κοινού συνάρτηση κατανομής είναι η συνάρτηση C . Πράγματι,

$$\begin{aligned} P[U_1 > u_1, U_2 > u_2] &= 1 - P[U_1 \leq u_1] + P[U_2 \leq u_2] + P[U_1 \leq u_1, U_2 \leq u_2] \\ &= 1 - u_1 - u_2 + C(u_1, u_2) \neq \bar{C}(u_1, u_2) \end{aligned}$$

αλλά,

$$P[U_1 > u_1, U_2 > u_2] = \bar{C}(1 - u_1, 1 - u_2)$$

Εφόσον η \bar{C} είναι συνάρτηση σύζευξης, παρεμβάλλεται μεταξύ του άνω και του κάτω ορίου Fréchet της C . Επομένως, ισχύει η παρακάτω ανισότητα για κάθε για $u \in [0, 1]^2$

$$C_L(u_1, u_2) \leq \bar{C}(u_1, u_2) \leq C_U(u_1, u_2)$$

Επιπλέον, μπορεί να επαληθευτεί ότι η ανεξάρτητη copula και το άνω και κάτω όριο Fréchet της C συμπίπτουν για τις copulas και τις copulas επιβίωσης. Δηλαδή,

$$\bar{C}_L = C_L, \quad \bar{C}_I = C_I \quad \text{και} \quad \bar{C}_U = C_U$$

Επίσης, να σημειώσουμε ότι το θεώρημα του Sklar μπορεί να επαναδιατυπωθεί ως προς τις copulas επιβίωσης. Συγκεκριμένα, η από κοινού συνάρτηση επιβίωσης \bar{F}_X του X μπορεί να εκφραστεί μέσω των περιθωρίων συναρτήσεων επιβίωσης \bar{F}_1 και \bar{F}_2 ($\bar{F} = 1 - F$) με την χρήση της copula επιβίωσης. Αυτό επιτυγχάνεται ως εξής

$$\bar{F}_X(x_1, x_2) = 1 - F_1(x_1) - F_2(x_2) + F_X(x) = \bar{C}(\bar{F}_1(x_1), \bar{F}_2(x_2))$$

Επομένως, η συνάρτηση \bar{C} συνδέει τις μονοδιάστατες περιθώριες \bar{F}_1 και \bar{F}_2 με την από κοινού \bar{F}_X με τρόπο ανάλογο με αυτόν που η C συνδέει την από κοινού σ.κ. F_X με τις περιθώριες της F_1 και F_2 .

2.3.2 Dual και co-copulas

Σε κάθε copula C αντιστοιχεί η co-copula C^* και η δυική (dual) συνάρτηση σύζευξης \tilde{C} που ορίζονται ως εξής:

Ορισμός: Η co-copula C^* , η οποία σχετίζεται με την copula C δίνεται από την σχέση

$$C^*(u_1, u_2) = 1 - C(1 - u_1, 1 - u_2), \quad u \in [0, 1]^2$$

ενώ η Dual \tilde{C} της copula C ισούται με

$$\tilde{C}(u_1, u_2) = u_1 + u_2 - C(u_1, u_2), \quad u \in [0, 1]^2$$

Να σημειώσουμε ότι οι co-copula C^* και dual \tilde{C} δεν αποτελούν συναρτήσεις copula, ωστόσο δίνονται από τις παρακάτω πιθανότητες

$$P[(X_1 > x_1) \cup (X_2 > x_2)] = C^*(\bar{F}_1(x_1), \bar{F}_2(x_2))$$

$$P[(X_1 \leq x_1) \cup (X_2 \leq x_2)] = \tilde{C}(F_1(x_1), F_2(x_2))$$

ενώ για την co-copula C^* ισχύει και η ακόλουθη ιδιότητα

$$(C^*)^* = C$$

Στις παρακάτω σχέσεις αποτυπώνεται η σύνδεση των co-copula C^* με τις ανεξάρτητες copulas και το άνω και κάτω όριο Fréchet της C ως εξής

$$C_L^*(u_1, u_2) = \min\{u_1 + u_2, 1\}$$

$$C_I^*(u_1, u_2) = u_1 + u_2 - u_1 u_2$$

$$C_U^*(u_1, u_2) = \max\{u_1, u_2\}$$

2.3.3 Αναλλοίωτοι μονότονοι μετασχηματισμοί (Functional Invariance)

Στην συγκεκριμένη παράγραφο θα παρουσιάσουμε την συμπεριφορά των copulas σε γνησίως μονότονους μετασχηματισμούς των τυχαίων μεταβλητών X_1 και X_2 .

Πρόταση: Έστω X_1 και X_2 συνεχείς τυχαίες μεταβλητές με copula C και t_1 και t_2 συνεχείς μονότονες συναρτήσεις.

- i. Εάν οι t_1 και t_2 είναι γνησίως αύξουσες τότε η copula του $(t_1(X_1), t_2(X_2))$ ισούται με την C .
- ii. Εάν η t_1 είναι γνησίως αύξουσα και η t_2 γνησίως φθίνουσα τότε η copula του $(t_1(X_1), t_2(X_2))$ ισούται με την $u_1 - C(u_1, 1 - u_2)$.
- iii. Εάν η t_1 είναι γνησίως φθίνουσα και η t_2 γνησίως αύξουσα τότε η copula του $(t_1(X_1), t_2(X_2))$ ισούται με την $u_2 - C(u_2, 1 - u_1)$.
- iv. Εάν οι t_1 και t_2 είναι γνησίως φθίνουσες τότε η copula του $(t_1(X_1), t_2(X_2))$ ισούται με την \bar{C} .

Επομένως, για γνησίως μονότονους μετασχηματισμούς των τυχαίων μεταβλητών η copula παραμένει αμετάβλητη ή υπολογίζεται με απλό τρόπο.

Σε κάθε μία από τις παραπάνω περιπτώσεις (i) - (iv), η μορφή της copula για $(t_1(X_1), t_2(X_2))$ είναι ανεξάρτητη από την επιλογή των t_1 και t_2 . Αυτό υποδεικνύει ότι η copula αποτυπώνει όλη την δομή εξάρτησης μεταξύ δύο συνεχών τυχαίων μεταβλητών, ο τρόπος με τον οποίο αλληλεπιδρούν οι X_1, X_2 αποδίδεται από την copula ανεξαρτήτως της κλίμακας στην οποία μετρώνται οι μεταβλητές.

Ας υποθέσουμε ότι έχουμε ένα πολυδιάστατο μοντέλο και τα δεδομένα μας αποτελούνται από εξαρτημένες ζημιές διαφόρων ειδών (πολυδιάστατη κατανομή). Εάν θελήσουμε να εξετάσουμε τον λογάριθμο των ζημιών αυτών, η συνάρτηση copula θα παραμείνει αμετάβλητη, ενώ οι περιθώριες κατανομές μπορεί να αλλάξουν.

Από την συμπεριφορά των συζεύξεων ως προς τους μονότονους μετασχηματισμούς, δηλαδή την ιδιότητα να παραμένει η copula αναλλοίωτη ενώ οι περιθώριες συναρτήσεις μπορούν να μεταβάλλονται, προκύπτουν επακόλουθα. Συγκεκριμένα, κάθε ιδιότητα της από κοινού συνάρτησης

κατανομής δύο τυχαίων μεταβλητών η οποία παραμένει αμετάβλητη από τον γνησίως αύξοντα μετασχηματισμό των τυχαίων μεταβλητών αποτελεί ιδιότητα και της copula συνάρτησης.

2.3.4 Εξάρτηση Ουράς

Στην αναλογιστική επιστήμη, η απόκλιση από την κανονικότητα οφείλεται συνήθως στις βαριές ουρές των δεδομένων. Σε ένα μοντέλο, οι βαριές ουρές μπορεί να προέρχονται από τις περιθώριες μονοδιάστατες κατανομές των τυχαίων μεταβλητών αλλά και από την από κοινού πιθανότητα εμφάνισης ακραίων τιμών.

Ας θεωρήσουμε ότι ένα συμβάν με μικρή πιθανότητα εμφάνισης παρατηρείται στην πρώτη μεταβλητή τυχαίου διανύσματος, δεδομένου ότι ένα συμβάν με την ίδια πιθανότητα εμφανίζεται στην δεύτερη μεταβλητή. Η Εξάρτηση Ουράς σχετίζεται ουσιαστικά με την πιθανότητα εμφάνισης μιας μεγάλης ζημίας για ένα συμβόλαιο δεδομένου ότι έχει συμβεί μια μεγάλη ζημιά σε ένα άλλο συμβόλαιο.

Η εξάρτηση Ουράς είναι σημαντική έννοια καθώς αποτελεί μέτρο εξάρτησης ακραίων τιμών των τυχαίων μεταβλητών. Ο συντελεστής εξάρτησης της άνω ουράς για ένα τυχαίο διάνυσμα (X_1, X_2) και με συνάρτηση copula C συμβολίζεται με λ_U και ισούται με

$$\begin{aligned}\lambda_U &= \lim_{v \rightarrow 0} P[X_1 > \bar{F}_1^{-1}(v) \mid X_2 > \bar{F}_2^{-1}(v)] \\ &= \lim_{v \rightarrow 0} \frac{P[X_1 > \bar{F}_1^{-1}(v), X_2 > \bar{F}_2^{-1}(v)]}{P[X_2 > \bar{F}_2^{-1}(v)]}\end{aligned}$$

ο αριθμητής της παραπάνω σχέσης μπορεί να γραφτεί ως εξής (για F_1, F_2 συνεχείς)

$$\begin{aligned}P[X_1 > \bar{F}_1^{-1}(v), X_2 > \bar{F}_2^{-1}(v)] &= 1 - P[X_1 \leq \bar{F}_1^{-1}(v)] - P[X_2 \leq \bar{F}_2^{-1}(v)] \\ &\quad + P[X_1 \leq \bar{F}_1^{-1}(v), X_2 \leq \bar{F}_2^{-1}(v)]\end{aligned}$$

επομένως, ο συντελεστής λ_U μπορεί να πάρει την μορφή

$$\lambda_U = \lim_{v \rightarrow 0} \frac{1 - 2(1 - v) + C(1 - v, 1 - v)}{v}$$

και μας οδηγεί στον ακόλουθο ορισμό (Joe 1993).

Ορισμός: Ο συντελεστής εξάρτησης ουράς λ_U ενός τυχαίου διανύσματος (X_1, X_2) με συνάρτηση Copula C ορίζεται ως εξής

$$\lambda_U = \lim_{v \rightarrow 1} \frac{1 - 2v + C(v, v)}{1 - v}$$

Αν $\lambda_U > 0$, X_1 και X_2 είναι ασυμπτωτικά εξαρτημένες, ενώ όταν $\lambda_U = 0$, οι τ.μ. X_1 και X_2 είναι ασυμπτωτικά ανεξάρτητες. Αυτό υποδεικνύει ότι όσο εξαρτημένες και να είναι οι τυχαίες μεταβλητές X_1 και X_2 , η εμφάνιση μίας ακραίας τιμής στην μία δεν εξαρτάται από την εμφάνιση μιας ακραίας τιμής στην άλλη μεταβλητή.

Να σημειώσουμε ότι ο συντελεστής εξάρτησης της κάτω ουράς προκύπτει με παρόμοιο τρόπο ως

$$\lambda_L = \lim_{v \rightarrow 0} \frac{C(v, v)}{v}$$

2.4 Αρχιμήδεια συνάρτηση σύζευξης (Archimedean Family of Copulas)

Οι Genest και MacKay (1986) περιέγραψαν μια σημαντική οικογένεια συναρτήσεων σύζευξης, την Αρχιμήδεια σύζευξη και παρουσίασαν εφαρμογές αυτής στην Στατιστική. Στην Αναλογιστική επιστήμη, οι περισσότερες εφαρμογές των συναρτήσεων σύζευξης επικεντρώθηκαν στην συγκεκριμένη οικογένεια συναρτήσεων (Frees και Valdez (1998), Denuit, Purcaru και Van Keilegom (2004)).

Ας δούμε ποια τα χαρακτηριστικά και πως ορίζεται η Αρχιμήδεια σύζευξη.

Ορισμός: Έστω συνάρτηση $\phi: [0, 1] \rightarrow \mathbb{R}^+$, με συνεχείς παραγώγους πρώτης και δεύτερης τάξης στο $(0, 1)$, τέτοια ώστε

$$\phi(1) = 0, \quad \phi'(\tau) < 0 \text{ και } \phi''(\tau) > 0 \quad \forall \tau \in (0, 1) \quad (3)$$

κάθε συνάρτηση ϕ που ικανοποιεί την (3) παράγει μία διδιάστατη συνάρτηση σύζευξη η οποία ορίζεται ως εξής

$$C_\phi(u_1, u_2) = \begin{cases} \phi^{-1}(\phi(u_1) + \phi(u_2)), & \phi(u_1) + \phi(u_2) \leq \phi(0) \\ 0, & \text{διαφορετικά} \end{cases} \quad (4)$$

όπου $0 \leq u_1$ και $u_2 \leq 1$.

Η Copula C_ϕ ονομάζεται Αρχιμήδεια σύζευξη ενώ η ϕ , η οποία ικανοποιεί την (3), καλείται γεννήτρια συνάρτηση της συνάρτησης σύζευξης. Για κάθε θετική σταθερά c , οι γεννήτριες ϕ και $c\phi$ παράγουν την ίδια σύζευξη. Επιπλέον, οι συνθήκες (3) αρκούν για να εγγυηθούν την ύπαρξη της ϕ^{-1} , αντίστροφη συνάρτηση της ϕ η οποία έχει επίσης δύο παραγώγους.

Επομένως, μια διδιάστατη σ.κ. $F_X \in \mathcal{R}_2(F_1, F_2)$ προσδιορίζεται μέσω μίας Αρχιμήδειας σύζευξης αν και μόνο αν μπορεί να εκφραστεί στην μορφή (1) (βλέπε Θεώρημα Sklar) για κάποια συνάρτηση $C = C_\phi$ η οποία ικανοποιεί τις (3) και (4).

Η οικογένεια της Αρχιμήδειας σύζευξης παρέχει μια σειρά μοντέλων τα οποία έχουν χρησιμοποιηθεί με επιτυχία στα πλαίσια μοντελοποίησης δεδομένων (Hennessy και Lapan (2002)). Επιπλέον, τα στοιχεία της συγκεκριμένης οικογένειας έχουν στοχαστικές ιδιότητες με αποτέλεσμα οι συναρτήσεις εξάρτησης C_ϕ είναι επιθυμητές για την στατιστική επεξεργασία δεδομένων.

Ένα παράδειγμα Αρχιμήδειας σύζευξης αποτελεί η οικογένεια Clayton με γεννήτρια συνάρτηση

$$\phi(t) = \frac{t^{-a} - 1}{a}$$

ομοίως, η σύζευξη Frank ανήκει στην Αρχιμήδεια οικογένεια με γεννήτρια συνάρτηση

$$\phi(t) = -\ln\left(\frac{\exp(-at) - 1}{\exp(-a) - 1}\right)$$

2.4.1 Σύνδεση Αρχιμήδειας Copula και Συναρτήσεων Πυκνότητας Πιθανότητας

Η συνάρτηση πυκνότητας πιθανότητας c_ϕ που σχετίζεται με την C_ϕ προκύπτει σύμφωνα με την ακόλουθα βήματα. Αρχικά, από την (4) προκύπτει ότι

$$\phi\left(C_\phi(u_1, u_2)\right) = \phi(u_1) + \phi(u_2)$$

την οποία παραγωγίζουμε ως προς την u_1

$$\phi'\left(C_\phi(u_1, u_2)\right) \frac{\partial}{\partial u_1} C_\phi(u_1, u_2) = \phi'(u_1) \quad (5)$$

στην συνέχεια παραγωγίζουμε την (5) ως προς την u_2

$$\phi''(C_\phi(u_1, u_2)) \frac{\partial}{\partial u_2} C_\phi(u_1, u_2) \frac{\partial}{\partial u_1} C_\phi(u_1, u_2) + \phi'(C_\phi(u_1, u_2)) \frac{\partial^2}{\partial u_1 \partial u_2} C_\phi(u_1, u_2) = 0$$

και αντικαθιστώντας στην παραπάνω σχέση τα στοιχεία $\frac{\partial}{\partial u_1} C_\phi(u_1, u_2)$ και $\frac{\partial}{\partial u_2} C_\phi(u_1, u_2)$ από την σχέση (5), δηλαδή

$$\frac{\partial}{\partial u_1} C_\phi(u_1, u_2) = \frac{\phi'(u_1)}{\phi'(C_\phi(u_1, u_2))}, \quad \frac{\partial}{\partial u_2} C_\phi(u_1, u_2) = \frac{\phi'(u_2)}{\phi'(C_\phi(u_1, u_2))}$$

προκύπτει η σ.π.π. $c_\phi(u_1, u_2)$ ως εξής:

$$c_\phi(u_1, u_2) = \frac{\partial^2}{\partial u_1 \partial u_2} C_\phi(u_1, u_2) = - \frac{\phi''(C_\phi(u_1, u_2)) \phi'(u_1) \phi'(u_2)}{(\phi'(C_\phi(u_1, u_2)))^3}$$

Από τις συνθήκες (3) προκύπτει ότι η $C_\phi(u) > 0$ για κάθε u τέτοιο ώστε

$$\phi(u_1) + \phi(u_2) < \phi(0)$$

Επιπλέον, να σημειώσουμε ότι οι παράγωγοι δεν υπάρχουν στο όριο $\phi(u_1) + \phi(u_2) = \phi(0)$.

Από τη σχέση (5) προκύπτει το κριτήριο Abel, μια χρήσιμη και επαρκή συνθήκη η οποία καθορίζει εάν μια συνάρτηση σύζευξης C είναι Αρχιμήδεια copula. Πιο συγκεκριμένα:

Μία συνάρτηση σύζευξης C είναι Αρχιμήδεια αν ορίζονται οι μερικές παράγωγοι $\frac{\partial}{\partial u_1} C(u_1, u_2)$ και $\frac{\partial}{\partial u_2} C(u_1, u_2)$ και αν υπάρχει συνάρτηση $g : (0, 1) \rightarrow (0, +\infty)$ τέτοια ώστε

$$g(u_2) \frac{\partial}{\partial u_1} C(u_1, u_2) = g(u_1) \frac{\partial}{\partial u_2} C(u_1, u_2), \quad \forall (u_1, u_2) \in [0, 1]^2$$

με αντίστοιχη γεννήτρια συνάρτηση

$$\phi(t) = \int_t^1 g(\xi) d\xi, \quad t \in [0, 1]$$

2.4.2 Ιδιότητες της Αρχιμήδειας σύζευξης

Στήριγμα

Η Αρχιμήδεια συνάρτηση σύζευξης C_ϕ έχει πεδίο ορισμού το συγκεκριμένο σύνολο

$$\{u \in [0, 1]^2 \mid \phi(u_1) + \phi(u_2) \leq \phi(0)\}$$

το οποίο ισούται με την τετραγωνική μονάδα $[0, 1]^2$ αν $\lim_{t \rightarrow 0^+} \phi(t) = +\infty$.

Ανεξαρτησία

Η σύζευξη C_I ανήκει στην Αρχιμήδεια οικογένεια. Πιο συγκεκριμένα, η C_I είναι Αρχιμήδεια ($C_\phi \equiv C_I$) αν και μόνο αν $\phi(t) = -c \ln t$, όπου c σταθερά $c > 0$.

Πράγματι, από την ισότητα $C_\phi = C_I$ προκύπτει η σχέση

$$\phi(u_1) + \phi(u_2) = \phi(u_1 u_2)$$

την οποία παραγωγίζοντας την ως προς u_1 και u_2 καταλήγουμε στην διαφορική εξίσωση

$$0 = \phi'(u_1 u_2) + u_1 u_2 \phi''(u_1 u_2).$$

Αν θέσουμε $u_1 u_2 = t$ η διαφορική εξίσωση μετατρέπεται ως εξής

$$\phi'(t) = -t \phi''(t)$$

της οποίας η λύση είναι η ζητούμενη σχέση $\phi(t) = -c \ln t$, $\phi(1) = 0$.

Ιδιάζοντα σύνολα της Αρχιμήδειας συνάρτησης σύζευξης

Υπάρχουν οικογένειες, οι οποίες ανήκουν στην Αρχιμήδεια σύζευξη και δεν είναι απόλυτα συνεχείς καθώς παρουσιάζουν γνήσια θετική πιθανότητα σε υποσύνολα του $[0,1]^2$ που έχουν μηδενικό μέτρο πιθανότητας Lebesgue. Σύμφωνα με τους Genest και MacKay (1986) έχουμε ιδιάζον σημείο όταν:

Πρόταση: Η σύζευξη C_ϕ έχει ιδιάζον σύνολο αν και μόνο αν $\frac{\phi(0)}{\phi'(0)} \neq 0$. Τότε,

$$\phi(U_1) + \phi(U_2) = \phi(0) \text{ με πιθανότητα } -\frac{\phi(0)}{\phi'(0)}$$

Αν $\frac{\phi(0)}{\phi'(0)} \neq 0$, η copula C_ϕ έχει θετική μάζα πιθανότητας πάνω στην καμπύλη $\phi(u_1) + \phi(u_2) = \phi(0)$. Επομένως, εάν η Αρχιμήδεια C_ϕ έχει ιδιάζον σημείο τότε αυτό περιέχεται στο σύνολο

$$\{u \in [0, 1]^2 | \phi(u_1) + \phi(u_2) = \phi(0)\}$$

και οι μερικές παράγωγοι $\frac{\partial}{\partial u_1} C_\phi$ και $\frac{\partial}{\partial u_2} C_\phi$ υπάρχουν παντού εκτός από την συγκεκριμένη καμπύλη (οριακή καμπύλη).

Επιπλέον όταν η ϕ' παρουσιάζει άλματα ασυνέχειας, μπορούν να προκύψουν άλλες ιδιαιτερότητες κατά την έκταση της καμπύλη $\phi(u_1) + \phi(u_2) = \phi(t)$ για τιμές t που αντιστοιχούν στα άλματα της ϕ' .

Για περισσότερες λεπτομέρειες για την οικογένεια της Αρχιμήδειας copula ο ενδιαφερόμενος αναγνώστης μπορεί να συμβουλευτεί τα συγγράμματα M. Denuit, J. Dhaene, M. Goovaerts, R. Kaas (2005) και Alexander J. McNeil, Rudiger Frey, Paul Embrechts (2005).

Κεφάλαιο 3

Μελέτη Ασφαλιστικού χαρτοφυλακίου

3.1 Εφαρμογή σε ασφαλιστικά δεδομένα ζημιών

Διάφορες διαδικασίες στην ασφάλιση ζημιών περιέχουν επεξεργασία συσχετιζόμενων στοιχείων, ένα χαρακτηριστικό παράδειγμα αποτελούν οι σωματικές βλάβες και οι υλικές ζημιές ενός ατυχήματος. Η αγνόηση πιθανής εξάρτησης μεταξύ των δεδομένων μπορεί να οδηγήσει σε υποεκτίμηση της ζημίας.

3.1.1 Δεδομένα

Σε αυτό το κεφάλαιο θα ασχοληθούμε με την πρακτική προσαρμογή σύζευξης σε ασφαλιστικά δεδομένα ζημιών και τον υπολογισμό της αντίστοιχης καμπύλης του VaR (Αξία σε Κίνδυνο) σε ποσοστό 99,5%. Για την υλοποίηση της εφαρμογής βασιζόμαστε στην γλώσσα προγραμματισμού R και τα δεδομένα που επεξεργαζόμαστε αποτελούνται από κλειστές ζημιές αστικής ευθύνης αυτοκινήτου. Επιπλέον, η εφαρμογή πραγματοποιείται σε διδιάστατα δεδομένα και για κάθε ζημιά εξετάζουμε τις εξής δύο μεταβλητές (BI_0 , PD_0):

- **το κόστος σωματικών βλαβών (BI_0 , Bodily Injuries)** που έχουν προκληθεί από το ασφαλιζόμενο όχημα (με υπαιτιότητα) σε άτομα όπως πεζούς, επιβαίνοντες άλλων οχημάτων που εμπλέκονται στο περιστατικό ή επιβαίνοντες του ίδιου οχήματος

- **το κόστος υλικών ζημιών (PD_0 , Property Damages)** που έχει προκαλέσει το ασφαλιζόμενο όχημα (με υπαιτιότητα) σε άλλα οχήματα ή άλλα αντικείμενα, όπως μάντρες, βιτρίνες, εισόδους καταστημάτων κ.α..

Το εξεταζόμενο δείγμα περιορίστηκε στις 4.800 ζημιές. Εξαιρέθηκαν οι ζημιές με μηδενικό κόστος, καθώς και τα συμβάντα με μικρά ποσά τα οποία αφορούσαν κυρίως έξοδα άμεσα συσχετιζόμενα με το ατύχημα, όπως αμοιβές πραγματογνωμόνων και δικηγόρων. Επιπλέον για την καλύτερη και πιο αποτελεσματική επεξεργασία, η ανάλυση των στοιχείων έγινε σε λογαριθμική κλίμακα. Επομένως, στους παρακάτω πίνακες παρουσιάζονται ορισμένα στατιστικά στοιχεία

του αρχικού δείγματος (σε ευρώ) και του αντίστοιχου δείγματος σε λογαριθμική κλίμακα (δηλαδή οι λογάριθμοι των αρχικών μεταβλητών).

| | Υλικές Ζημιές (PD ₀) | Σωματικές Βλάβες (BI ₀) |
|-----------------|----------------------------------|-------------------------------------|
| MAX | 282,674 | 737,365 |
| MIN | 505 | 1,956 |
| SD | 9,876 | 63,742 |
| Average | 4,442 | 25,466 |
| Median | 2,134 | 7,046 |
| 25th Percentile | 1,064 | 3,497 |
| 75th Percentile | 4,750 | 17,829 |

Στατιστικά στοιχεία αρχικού δείγματος.

| | ln(PD ₀) | ln(BI ₀) |
|-----------------|----------------------|----------------------|
| MAX | 12.55 | 13.51 |
| MIN | 6.22 | 7.58 |
| SD | 0.99 | 1.21 |
| Average | 7.78 | 9.12 |
| Median | 7.67 | 8.86 |
| 25th Percentile | 6.97 | 8.16 |
| 75th Percentile | 8.47 | 9.79 |

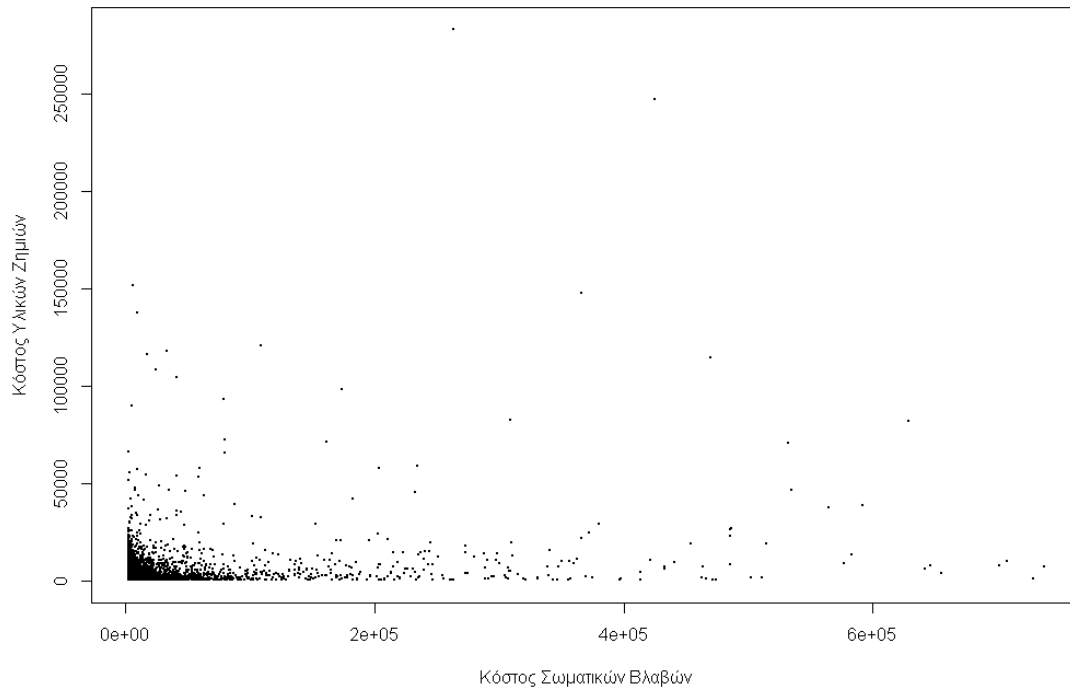
Στατιστικά στοιχεία δείγματος σε λογαριθμική κλίμακα.

Στην συνέχεια του κεφαλαίου θα συμβολίζουμε τους λογαρίθμους των αρχικών τιμών, δηλαδή τις τιμές $\ln(BI_0)$, $\ln(PD_0)$, με BI και PD αντίστοιχα.

Σχηματίζοντας το γράφημα διασποράς των δύο μονοδιάστατων μεταβλητών BI, PD φαίνεται από την κατανομή των στοιχείων ότι οι μικρές τιμές της μεταβλητής BI τείνουν να σχετίζονται με μικρές τιμές της μεταβλητής PD.

```
> library("readxl")
> my_data <- read_excel("Sample_BI_PD.xlsx")
> plot(my_data, main="Scatterplot", xlab="Κόστος Σωματικών Βλαβών",
ylab = "Κόστος Υλικών Ζημιών", pch=20, cex=.1)
```

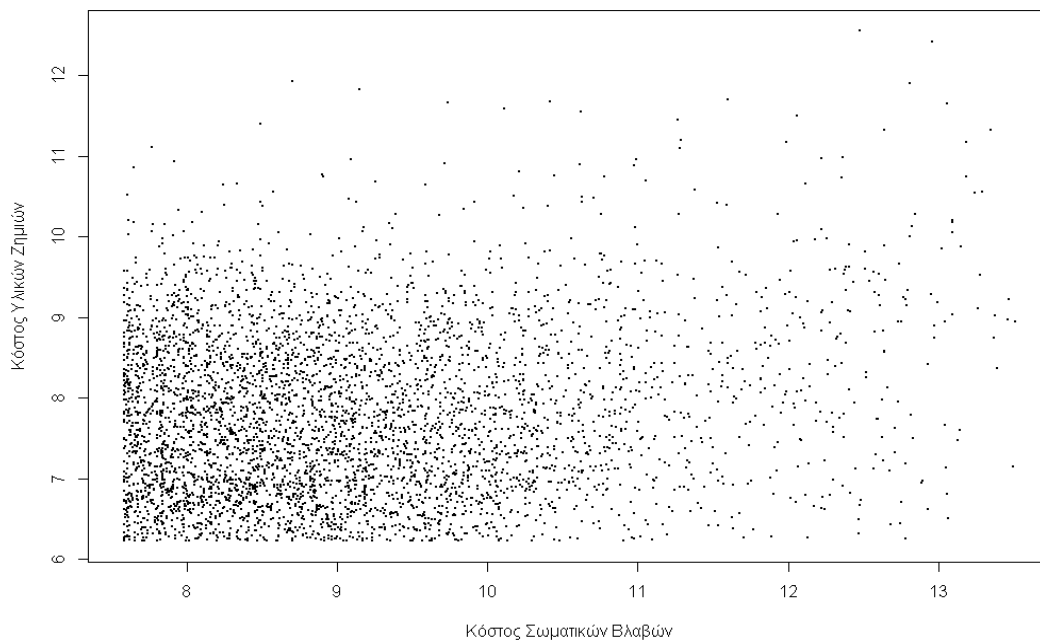
Scatterplot



```
> logtable<-log(my_data)
```

```
> plot(logtable, main="log Scatterplot", xlab="Κόστος Σωματικών Βλαβών", ylab="Κόστος Υλικών Ζημιών", pch=20, cex=.1)
```

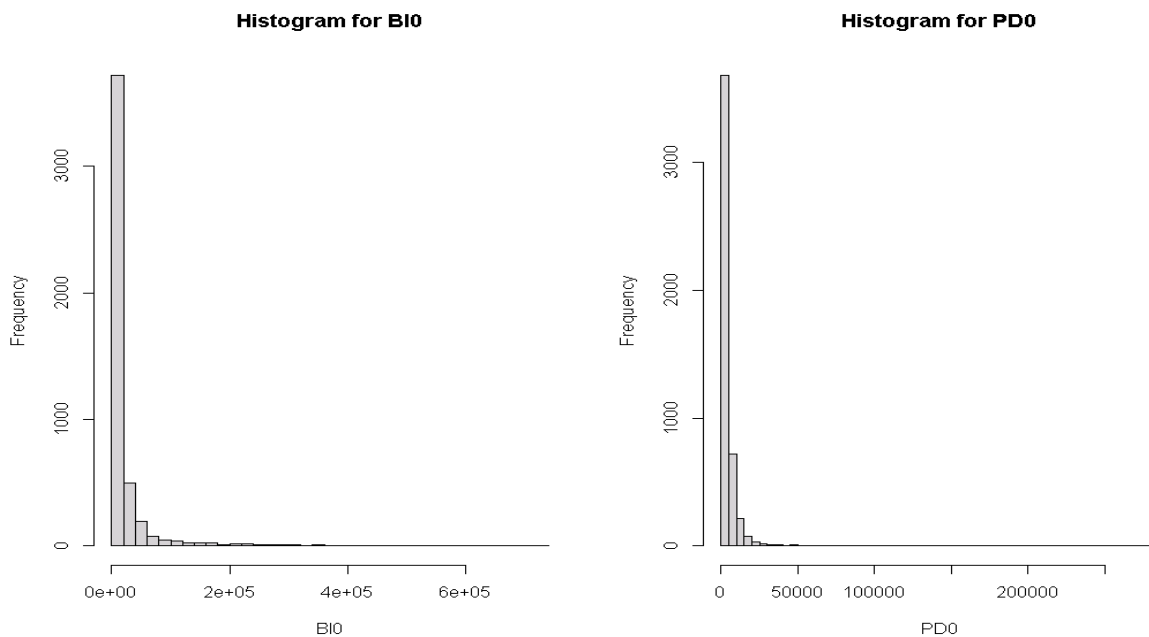
log Scatterplot



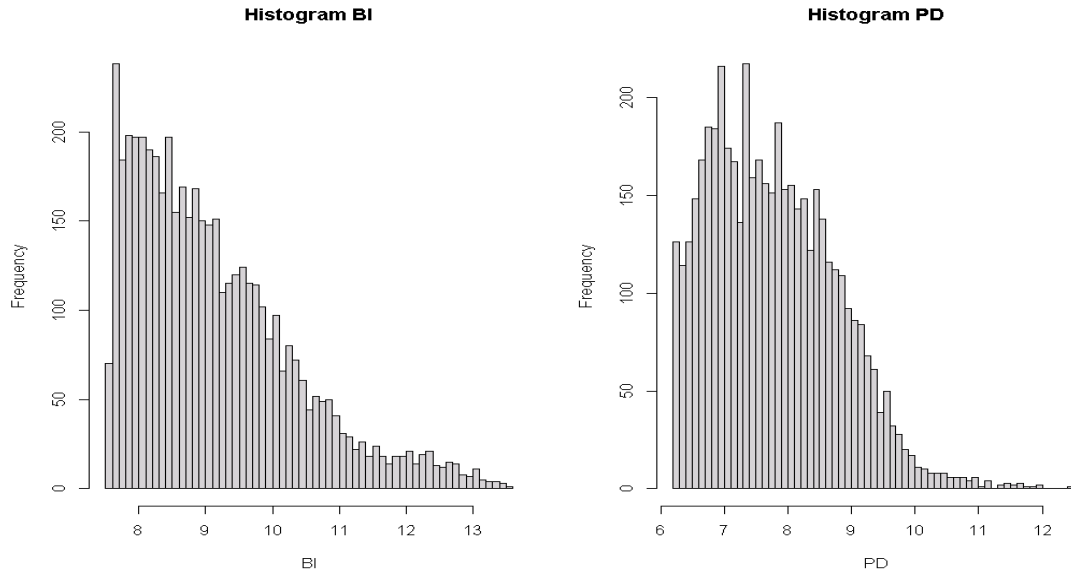
3.1.2 Περιθώριες κατανομές των μεταβλητών BI και PD

Για τον υπολογισμό της από κοινού συνάρτησης κατανομής χρειάζεται να προσδιορίσουμε αρχικά κατάλληλες περιθώριες κατανομές για τις μονοδιάστατες μεταβλητές BI και PD. Παρατηρώντας τα ιστογράμματα τα οποία κατασκευάσαμε για τις δύο μεταβλητές, και δεδομένου ότι τα στοιχεία τα οποία επεξεργαζόμαστε είναι σε λογαριθμική κλίμακα με μικρή ουρά, η κατανομή Γάμμα πιθανόν να προσαρμόζεται ικανοποιητικά στα δεδομένα των BI, PD. Εφόσον έχουμε περικόψει τα αρχικά δεδομένα μας εννοείται ότι θα πρέπει να βασιστούμε σε προσέγγιση από περικομμένες (από αριστερά) γάμμα κατανομές. Αυτό ίσως διαφαίνεται και από τα ακόλουθα ιστογράμματα των BI, PD.

```
> par(mfrow = c(1, 2))
> hist(BI0,main="Histogram for BI0", xlab = "BI0",50)
> hist(PD0,main="Histogram for PD0", xlab = "PD0",50)
```



```
> BI<-log(BI0);PD<-log(PD0)
> par(mfrow = c(1, 2))
> hist(BI,main="Histogram BI", xlab = "BI",50)
> hist(PD,main="Histogram PD", xlab = "PD",50)
```



Η υπόθεση ότι οι περιθώριες κατανομές των λογαρίθμων των δεδομένων μας ακολουθούν προσεγγιστικά μια γάμμα κατανομή υποδηλώνει ότι οι αρχικές τιμές προέρχονται από κατανομή με βαριά δεξιά ουρά. Αυτό γίνεται φανερό και από το ότι αν μια μεταβλητή X ακολουθεί εκθετική κατανομή, $X \sim \text{Exp}$, τότε η $Y = e^X$ ακολουθεί την κατανομή Pareto που ως γνωστό έχει βαριά δεξιά ουρά. Επομένως αν $X \sim \text{Gamma}(\kappa, \lambda)$ τότε η $Y = e^X$ θα γράφεται ως γινόμενο ανεξάρτητων Pareto κατανομών.

Το επόμενο βήμα είναι να εκτιμήσουμε κατάλληλες παραμέτρους για τις περιθώριες κατανομές ώστε να προσαρμόζονται καλά στα δεδομένα. Αυτό θα γίνει χρησιμοποιώντας το R και συγκεκριμένα το πακέτο `fitdistrplus` όπου θεωρείται ότι ένα ποσοστό των δεδομένων (pBI και pPD) κάτω από το `minvalue` είναι περικομμένα. Οι εκτιμήσεις των παραμέτρων μέσω της εντολής `fitdistcens` γίνεται αριθμητικά χρησιμοποιώντας την μέθοδο μέγιστης πιθανοφάνειας.

```
# εκτίμηση παραμέτρων περικομμένης γάμμα κατανομής για το BI
> library(fitdistrplus)
> n=length(BI)
> pBI=0.4;

#η τιμή στην οποία περικόπτουμε αριστερά τα δεδομένα μας
> minvalue=7.5; n1=n/(1-pBI);
> BIs=data.frame(left=c(BI, rep(NA, n1-n)), right=c(BI, rep(minvalue, n1-n)))
```

```

> a=fitdistcens(BIs, "gamma")
> shapeBI=a$estimate[1];scaleBI=a$estimate[2]
> hist(BI,100,prob=TRUE,col="gray")
> curve(dgamma(x,shapeBI,scaleBI)/(1-pBI),4,14,add=TRUE,col="blue")
> shapeBI;scaleBI
  shape
22.74349
  rate
2.824626

```

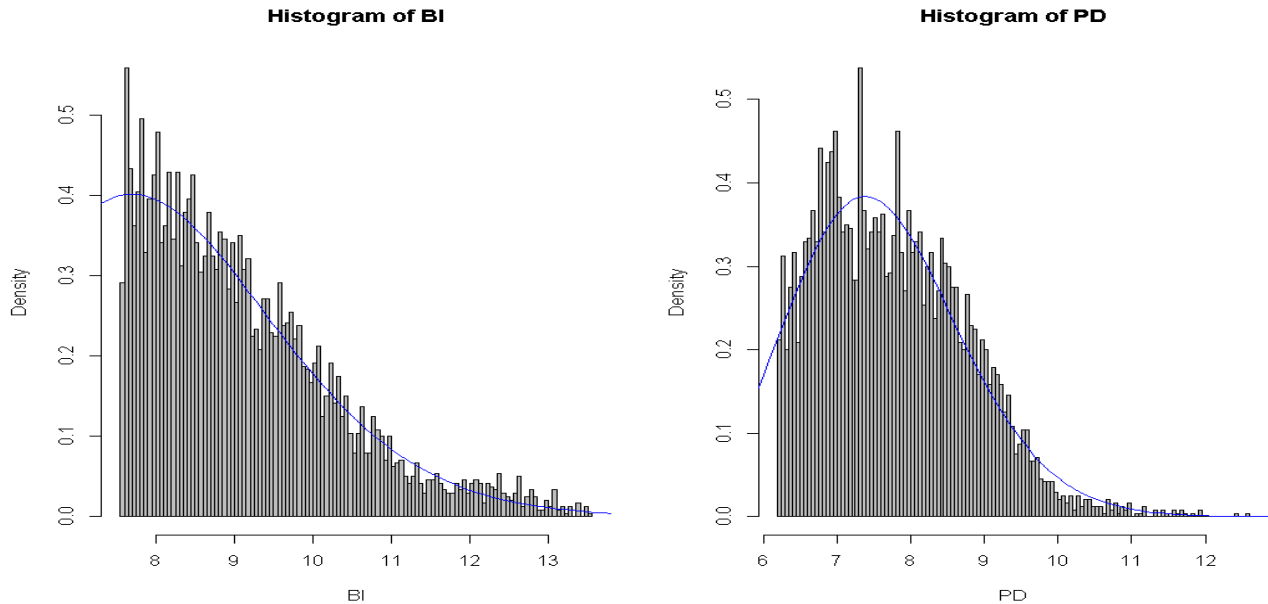
Επομένως, για τις τιμές της μεταβλητής BI εκτιμάται ότι προέρχονται προσεγγιστικά από την κατανομή Γάμμα (περικομμένη αριστερά στο 7.5) με παραμέτρους σχήματος και κλίμακας 22.7 και 2.82 αντίστοιχα.

```

# εκτίμηση παραμέτρων περικομμένης γάμμα κατανομής για το PD
> pPD=0.1;
#η τιμή στην οποία περικόπτουμε αριστερά τα δεδομένα μας
> minvalue=6; n1=n/(1-pPD);
> PDs=data.frame(left=c(PD,rep(NA,n1-n)),right=c(PD,rep(minvalue,n1-
n)))
> a=fitdistcens(PDs, "gamma")
> shapePD=a$estimate[1];scalePD=a$estimate[2]
> hist(PD,100,prob=TRUE,col="gray")
> curve(dgamma(x,shapePD,scalePD)/(1-pPD),4,14,add=TRUE,col="blue")
> shapePD;scalePD
  shape
42.05651
  rate
5.562887

```

Όμοια για τις τιμές της μεταβλητής PD εκτιμάται ότι προέρχονται προσεγγιστικά από την κατανομή Γάμμα (περικομμένη αριστερά στο 6) με παραμέτρους σχήματος και κλίμακας 42.05 και 5.56 αντίστοιχα.



Στα παραπάνω ιστογράμματα των μεταβλητών BI και PD έχουμε πλέον προσαρμόσει τις συναρτήσεις πυκνότητας των (περικομμένων) κατανομών Γάμμα(22.74, 2.82) και Γάμμα(45.06, 5.6) αντίστοιχα οι οποίες φαίνεται να κάνουν καλή εφαρμογή στα δεδομένα.

Στο σημείο αυτό μπορούμε πλέον να ξεκινήσουμε την μελέτη της από κοινού συμπεριφοράς των μεταβλητών.

3.1.3 Προσαρμογή Σύζευξης σε διδιάστατα δεδομένα

Αφετηρία για την προσαρμογή σύζευξης σε δεδομένα αποτελεί ο προσδιορισμός κατάλληλης οικογένειας συζεύξεων. Πριν προχωρήσουμε όμως σε αυτό το στάδιο, θα υπολογίσουμε τον συντελεστή συσχέτισης Kendall τ (tau).

Ο συντελεστής συσχέτισης Kendall tau αποτελεί ένα μη παραμετρικό μέτρο εξάρτησης και ο υπολογισμός του βασίζεται στην σχετική θέση των παρατηρήσεων και όχι από τις ίδιες τις τιμές τους, έτσι ώστε να μην επηρεάζεται από μονότονους μετασχηματισμούς των παρατηρήσεων. Συγκεκριμένα, η εκτίμηση του συντελεστή συσχέτισης εξαρτάται από το πλήθος των εναρμονισμένων ή συσχετισμένων (concordant) παρατηρήσεων και το πλήθος των μη εναρμονισμένων ή

μη συσχετισμένων (discordant) παρατηρήσεων, όπου οι παρατηρήσεις αποτελούν ζεύγη τιμών (X_1, X_2) .

Δύο παρατηρήσεις με αντίστοιχα ζεύγη τιμών (X_1, X_2) και (X'_1, X'_2) είναι εναρμονισμένες όταν και οι δύο τιμές X_1, X_2 της μίας παρατήρησης είναι μεγαλύτερες ή μικρότερες από τις αντίστοιχες τιμές της δεύτερης παρατήρησης. Αν ισχύει $X_1 < X'_1$ και $X_2 < X'_2$ ή αν $X_1 > X'_1$ και $X_2 > X'_2$ τότε οι παρατηρήσεις είναι εναρμονισμένες, ενώ αν ισχύει $X_1 < X'_1$ και $X_2 > X'_2$ ή $X_1 > X'_1$ και $X_2 < X'_2$ τότε οι παρατηρήσεις είναι μη εναρμονισμένες. Επιπλέον, οι (X_1, X_2) και (X'_1, X'_2) είναι εναρμονισμένες αν $(X_1 - X'_1)(X_2 - X'_2) > 0$, ενώ είναι μη εναρμονισμένες αν $(X_1 - X'_1)(X_2 - X'_2) < 0$.

Συνεπώς, ισχύει ότι:

$$P[\text{εναρμονισμένου ζεύγους}] = P[(X_1 - X'_1)(X_2 - X'_2) > 0]$$

και

$$P[\text{μη εναρμονισμένου ζεύγους}] = P[(X_1 - X'_1)(X_2 - X'_2) < 0]$$

Οι πιθανότητες γεγονότων που αφορούν σχέσεις ανισότητας μεταξύ δύο τυχαίων μεταβλητών, παραμένουν αμετάβλητες από τον γνησίως αύξοντα μετασχηματισμό των μεταβλητών αυτών. Επομένως, ο προσδιορισμός ενός μέτρου εξάρτησης μέσω των πιθανοτήτων πλήθους εναρμονισμένων ή μη εναρμονισμένων παρατηρήσεων διασφαλίζει την εξάρτηση του από την υποκείμενη σύζευξη.

Ορισμός: Ο συντελεστής συσχέτισης Kendall tau ενός τυχαίου ζεύγους (X_1, X_2) ορίζεται ως εξής

$$r_K(X_1, X_2) = P(\text{εναρμονισμένου ζεύγους}) - P(\text{μη εναρμονισμένου ζεύγους})$$

Ο αντίστοιχος δειγματικός συντελεστής συσχέτισης του Kendall που θα υπολογίσουμε στην συνέχεια είναι

$$\tau = \frac{\text{πλήθος εναρμονισμένων ζευγών} - \text{πλήθος μη εναρμονισμένων ζευγών}}{n(n-1)/2}$$

όπου n το πλήθος των παρατηρήσεων του δείγματος.

Αν οι περιθώριες των X_1, X_2 είναι συνεχείς κατανομές, τότε ο συντελεστής r_K ορίζεται ως

$$r_K(X_1, X_2) = 2P[(X_1 - X'_1)(X_2 - X'_2) > 0] - 1$$

όπου (X'_1, X'_2) είναι ανεξάρτητο αντίγραφο του (X_1, X_2) , δηλαδή είναι ανεξάρτητα και έχουν την ίδια από κοινού κατανομή.

Ο συντελεστής Kendall tau παίρνει τιμές στο διάστημα $[-1, 1]$ και παραμένει αμετάβλητος σε γνησίως μονότονους μετασχηματισμούς. Ισχύει ότι αν t_1, t_2 μη φθίνουσες συνεχείς συναρτήσεις τότε

$$r_K(t_1(X_1), t_2(X_2)) = r_K(X_1, X_2)$$

Με βάση την παραπάνω σχέση, αν οι X_1 και X_2 έχουν συνεχείς περιθώριες συναρτήσεις κατανομής F_1 και F_2 τότε

$$r_K(F_1(X_1), F_2(X_2)) = r_K(X_1, X_2)$$

συνεπώς, το μέτρο r_K εξαρτάται μόνο από την σύζευξη των (X_1, X_2) .

Αν (X_1, X_2) ζεύγος συνεχών τυχαίων μεταβλητών με σύζευξη C και (X'_1, X'_2) ανεξάρτητο αντίγραφο $(X_1, X_2) =_d (X'_1, X'_2)$, ο συντελεστής συσχέτισης r_K μπορεί να εκφραστεί στην μορφή

$$\begin{aligned} r_K(X_1, X_2) &= 2P[(X_1 - X'_1)(X_2 - X'_2) > 0] - 1 = 4P[X_1 \leq X'_1, X_2 \leq X'_2] - 1 \\ &= 4 \int_0^1 \int_0^1 C(u_1, u_2) dC(u_1, u_2) - 1 \end{aligned}$$

όπου (U_1, U_2) ζεύγος ομοιόμορφων $U(0, 1)$ τυχαίων μεταβλητών με από κοινού συνάρτηση κατανομής C .

Στην περίπτωση της Αρχιμήδειας σύζευξης C_ϕ , ο συντελεστής r_K υπολογίζεται άμεσα μέσω της γεννήτριας συνάρτησης ϕ και δίνεται από την σχέση

$$r_K^\phi = 4 \int_0^1 \frac{\phi(t)}{\phi'(t)} dt + 1$$

Επιστρέφοντας στην εφαρμογή, ο δειγματικός συντελεστής συσχέτισης Kendall tau των δύο μεταβλητών BI και PD ισούται με 0,045 υποδεικνύοντας ότι υπάρχει συσχέτιση μεταξύ των δύο μεταβλητών αλλά όχι έντονη.

```
> data<-data.frame(BI,PD)
> cor(data, method = "kendall")
          BI          PD
BI 1.00000000 0.04508195
PD 0.04508195 1.00000000
```

Στην συνέχεια και αφού έχοντας προσεγγίσει με την βοήθεια συγκεκριμένων εντολών της γλώσσας προγραμματισμού R τις περιθώριες κατανομές, θα προσδιορίσουμε την οικογένεια σύζευξης που προσαρμόζεται καλύτερα στα διδιάστατα δεδομένα μας. Οι παράμετροι κάθε οικογένειας σύζευξης υπολογίζονται με βάση την Μέθοδο Μέγιστης Πιθανοφάνειας και η επιλογή της καταλληλότερης οικογένειας γίνεται με βάση την μικρότερη τιμή του κριτηρίου AIC (Akaike Information Criterion).

Το κριτήριο AIC εκτιμάει την ποιότητα εφαρμογής ενός μοντέλου σε ένα σύνολο δεδομένων συγκριτικά με άλλα μοντέλα και χρησιμοποιείται ευρέως στον προσδιορισμό του καταλληλότερου. Ο υπολογισμός του AIC βασίζεται στον αριθμό των ανεξάρτητων μεταβλητών που χρησιμοποιούνται για την δημιουργία του μοντέλου και την εκτίμηση μέγιστης πιθανοφάνειας του μοντέλου. Σύμφωνα με το κριτήριο, το καταλληλότερο μοντέλο είναι εκείνο το οποίο προσεγγίζει την απόκλιση από τα δεδομένα χρησιμοποιώντας τις λιγότερες δυνατές ανεξάρτητες μεταβλητές.

Για την εύρεση της βέλτιστης συνάρτησης σύζευξης χρησιμοποιούμε το πακέτο VineCopula και την εντολή BiCopSelect.

```
> library(VineCopula)
> var_a<-pobs(BI)
> var_b<-pobs(PD)
> selectedcopula<-BiCopSelect(var_a, var_b, familyset=NA)
> selectedcopula
Bivariate copula: Joe (par = 1.12, tau = 0.07)
```

```
> summary(selectedcopula)
```

```
Family
```

```
-----
```

```
No:      6
```

```
Name:    Joe
```

```
Parameter(s)
```

```
-----
```

```
par:     1.12
```

```
Dependence measures
```

```
-----
```

```
Kendall's tau:    0.07
```

```
Upper TD:         0.15
```

```
Lower TD:         0
```

```
Fit statistics
```

```
-----
```

```
logLik:    68.35
```

```
AIC:       -134.7
```

```
BIC:       -128.23
```

Συνεπώς, σύμφωνα με την παραπάνω διαδικασία, η οικογένεια σύζευξης Joe αποτελεί μία καλή επιλογή για τα δεδομένα μας.

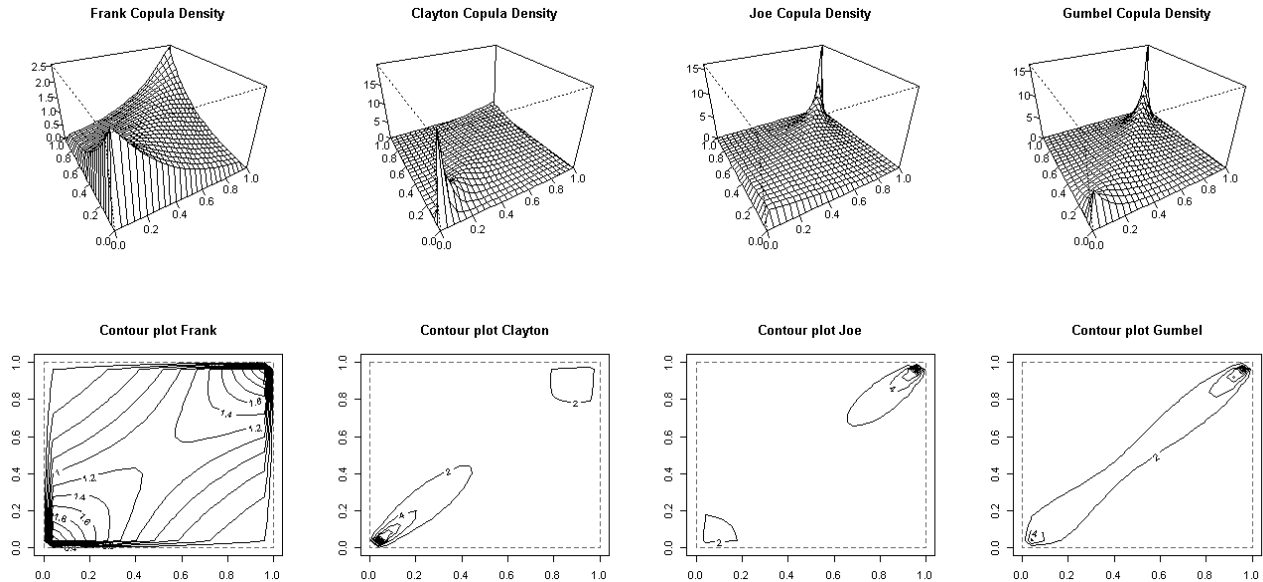
Η Joe σύζευξη ανήκει στην Αρχιμήδεια οικογένεια και παρουσιάστηκε από τον Joe (1993, 1997). Η συνάρτηση C δίνεται από τον τύπο:

$$C_a(u_1, u_2) = 1 - [(1 - u_1)^a + (1 - u_2)^a - (1 - u_1)^a(1 - u_2)^a]^{1/a}, \quad 1 \leq a < \infty$$

και έχει γεννήτρια συνάρτηση ϕ ,

$$\phi(t) = -\ln[1 - (1 - t)^\alpha]$$

Εάν, $\alpha = 1$ τότε καταλήγουμε στην ειδική περίπτωση της ανεξάρτητης σύζευξης, ενώ όταν $\alpha \rightarrow \infty$ προκύπτει το άνω όριο Fréchet της σύζευξης.



Στα παραπάνω γραφήματα παρίσταται η πυκνότητα για τις εξής οικογένειες της Αρχιμήδειας σύζευξης: Frank, Clayton, Joe και Gumbel. Οι παράμετροι α έχουν οριστεί αυθαίρετα.

Εφόσον προσδιορίσαμε την οικογένεια σύζευξης θα εκτιμήσουμε την αντίστοιχη παράμετρο α με την μέθοδο της μέγιστης πιθανοφάνειας. Για το σκοπό αυτό χρησιμοποιούμε την εντολή `fitCopula` από το πακέτο `copula`.

```
# εγκαθιστώ στο πακέτο με εντολές των συζεύξεων
> library(copula)
> copula<-joeCopula(dim=2)
> t<-pobs(as.matrix(logtable))
> fit<-fitCopula(copula,t, method='ml')
> coef(fit)
alpha
1.124687
> tau(joeCopula(param=as.numeric(coef(fit))))
```

```
[1] 0.06681795
```

Επομένως, καταλήγουμε στην Joe σύζευξη με παράμετρο 1.12. Παρατηρούμε ότι ο συντελεστής συσχέτισης Kendall tau της σύζευξης είναι αρκετά κοντά στον αντίστοιχο συντελεστή που υπολογίστηκε με βάση τις παρατηρήσεις BI, PD.

Με βάση το μοντέλο μας, η σύζευξη η οποία προσαρμόζεται καλύτερα στα δεδομένα μας είναι η Joe. Ωστόσο, χρησιμοποιώντας τις παραπάνω εντολές της R μπορούμε να εφαρμόσουμε διάφορες οικογένειες συζεύξεων στα δεδομένα και να παρατηρήσουμε τις αντίστοιχα γραφήματα που προκύπτουν. Για παράδειγμα, παρουσιάζουμε την εφαρμογή των οικογενειών Frank, Gumbel και Normal (Gaussian) στα δεδομένα εκτιμώντας τις αντίστοιχες παραμέτρους με την μέθοδο της μέγιστης πιθανοφάνειας και κατασκευάζοντας τις απεικονίσεις αυτών.

```
> copula_t1<-frankCopula(dim=2)
> t1<-pobs(as.matrix(logtable))
> fit_t1<-fitCopula(copula_t1,t1, method='ml')
> coef(fit_t1)
```

```
alpha
```

```
0.4113832
```

```
> copula_t2<-gumbelCopula(dim=2)
> t2<-pobs(as.matrix(logtable))
> fit_t2<-fitCopula(copula_t2,t2, method='ml')
> coef(fit_t2)
```

```
alpha
```

```
1.078317
```

```
> copula_t3<-normalCopula(dim=2)
> t3<-pobs(as.matrix(logtable))
> fit_t3<-fitCopula(copula_t3,t3, method='ml')
> coef(fit_t3)
```

```
rho.1
```

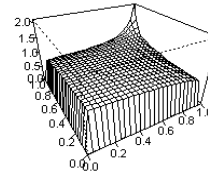
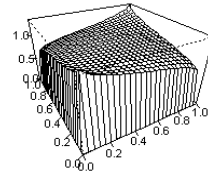
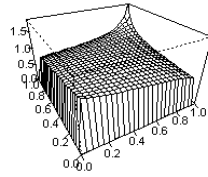
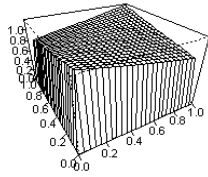
```
0.09918814
```

Αφού υπολογίσαμε την παράμετρο κάθε σύζευξης, κατασκευάζουμε στην συνέχεια τις απεικονίσεις αυτών.

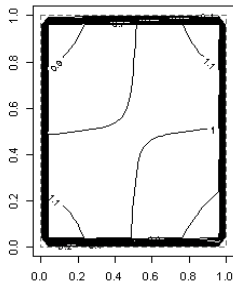
```
> par(mfrow = c(2, 4))
> frank<-frankCopula(dim=2, param=0.4113832)
> normal<-normalCopula(dim=2, param=0.09918814)
> joe<-joeCopula(dim=2, param=1.124687)
> gumbel<-gumbelCopula(dim=2, param=1.078317)

> persp(frank, dCopula, main="Frank Copula Density with param=0.41")
> persp(gumbel, dCopula, main="Gumbel Copula Density with param=1.08")
> persp(normal, dCopula, main="Normal Copula Density with
param=0.099")
> persp(joe, dCopula, main="Joe Copula Density with param=1.12")
> contour(frank, dCopula,xlim=c(0,1), ylim=c(0,1), main="Contour plot
Frank with param=0.41")
> contour(gumbel,dCopula,xlim=c(0,1), ylim=c(0,1), main="Contour plot
Gumbel with param=1.08")
> contour(normal, dCopula,xlim=c(0,1),ylim=c(0,1), main="Contour plot
Normal with param=0.099")
> contour(joe, dCopula,xlim=c(0,1), ylim=c(0,1), main="Contour plot
Joe with param=1.12")
```

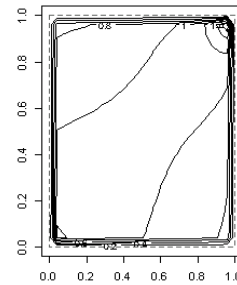
Frank Copula Density with param=0.41 Gumbel Copula Density with param=1.0 Normal Copula Density with param=0.09 Joe Copula Density with param=1.12



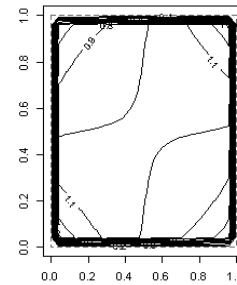
Contour plot Frank with param=0.41



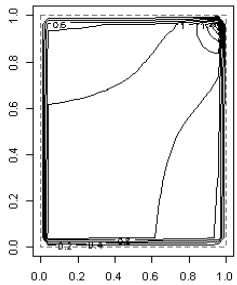
Contour plot Gumbel with param=1.08



Contour plot Normal with param=0.099



Contour plot Joe with param=1.12



3.1.4 Προσδιορισμός της Διδιάστατης κατανομής με την χρήση Σύζευξης

Εφόσον έχουμε προσδιορίσει την σύζευξη που προσαρμόζεται καλύτερα στα δεδομένα μας και έχουμε επιλέξει και τις αντίστοιχες περιθώριες κατανομές των BI, PD, μπορούμε να προσεγγίσουμε την από κοινού συμπεριφορά των περιθωρίων μεταβλητών.

#κατασκευή της διδιάστατης από κοινού κατανομής

```
> dist<-mvd(joeCopula(param=1.124687, dim=2), margins=c("gamma",
"gamma"), paramMargins=list(list(shape=shapeBI,rate=scaleBI),list(shape=shapePD,rate=scalePD)))
```

#παραγωγή τυχαίου δείγματος 5000 παρατηρήσεων από την από κοινού κατανομή

```
> v<-rMvdc(5000,dist)
```

#υπολογισμός συνάρτησης πυκνότητας

```
> pdf<-dMvdc(v,dist)
```

#υπολογισμός συνάρτησης κατανομής

```

> cdf<-pMvdc(v,dist)

> library(scatterplot3d)

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

> scatterplot3d(v[,1],v[,2], pdf, color="red", main="Density plot for
Bivariate distribution", xlab = "BI", ylab="PD", zlab="dMvdc",pch=".")

> scatterplot3d(v[,1],v[,2], cdf, color="red", main="Cumulative dis-
tribution function", xlab = "BI", ylab="PD", zlab="pMvdc",pch=".")

> persp(dist, dMvdc, xlim = c(0, 14), ylim=c(0, 14), main = "Density
plot for Bivariate distribution", xlab = "BI", ylab="PD")

> contour(dist, dMvdc, xlim = c(0, 14), ylim=c(0, 14), main = "Density
Contour plot",xlab = "BI", ylab="PD")

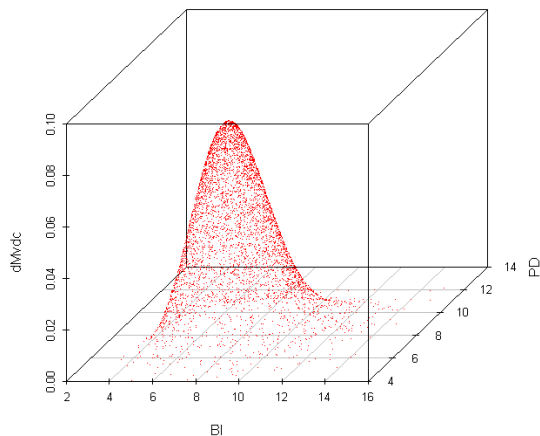
> persp(dist,pMvdc, xlim = c(0, 14), ylim=c(0, 14), main = "Cumulative
distribution function", xlab = "BI", ylab="PD")

> contour(dist,pMvdc, xlim = c(0, 15), ylim=c(0, 14),xlab = "BI",
ylab="PD", main = "Cumulative distribution Contour plot", lev-
els=c(seq(0.1,0.9, by=0.1),0.99,0.995,0.999))

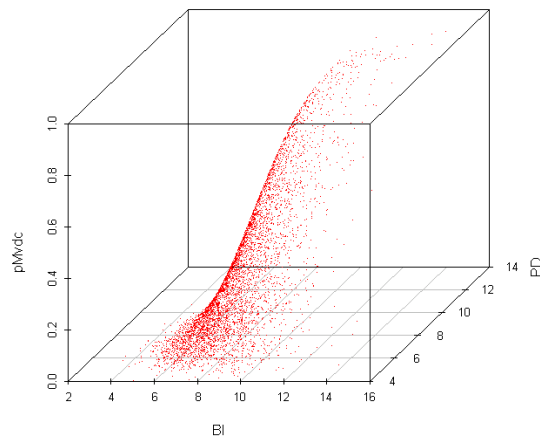
```

Με την βοήθεια της γλώσσας R, προσδιορίσαμε την από κοινού συνάρτηση μέσω της σύζευξης Joe με παράμετρο 1.12 και με περιθώριες Γάμμα(22.74, 2.82) και Γάμμα(45.06, 5.6). Για να δούμε την συμπεριφορά των παρατηρήσεων που επεξεργαζόμαστε κατασκευάσαμε ένα τυχαίο δείγμα 5000 παρατηρήσεων από την διδιάστατη από κοινού κατανομή και υπολογίσαμε την συνάρτηση πυκνότητας και την συνάρτηση κατανομής, τις απεικονίσεις των οποίων μπορούμε να παρατηρήσουμε στα παρακάτω γραφήματα μέσα από διαφορετικές οπτικές και μορφές.

Density plot for Bivariate distribution

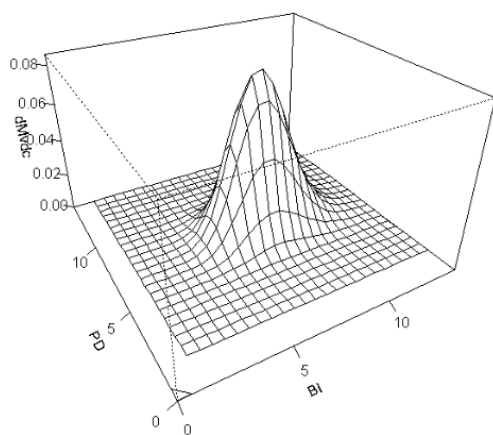


Cumulative distribution function

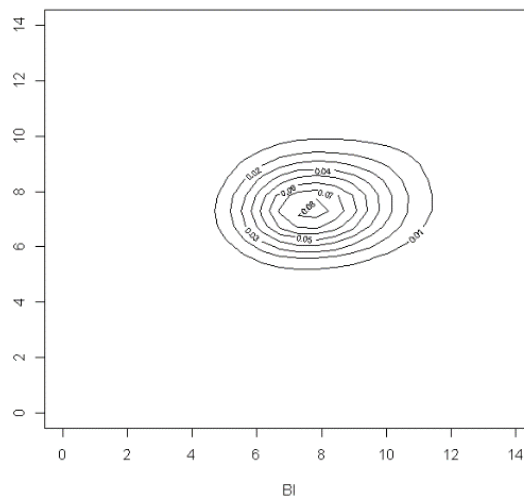


Η από κοινού συνάρτηση πυκνότητας πιθανότητας της κατανομής που προσαρμόζεται στα δεδομένα μας δίνεται στο ακόλουθο σχήμα (σύζευξη Joe με περιθώριες κατανομές Γάμμα). Συνεπώς, οι λογάριθμοι των τιμών των 2 μεταβλητών μας θα κατανέμονται γύρω από τις τιμές 8 και 7 περίπου όπως φαίνεται στο ακόλουθο δεξί σχήμα.

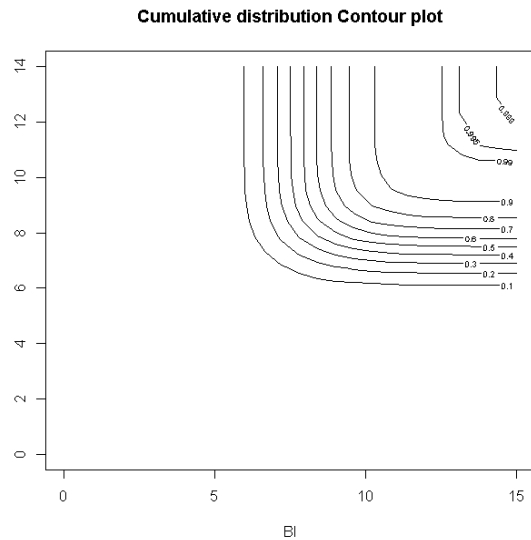
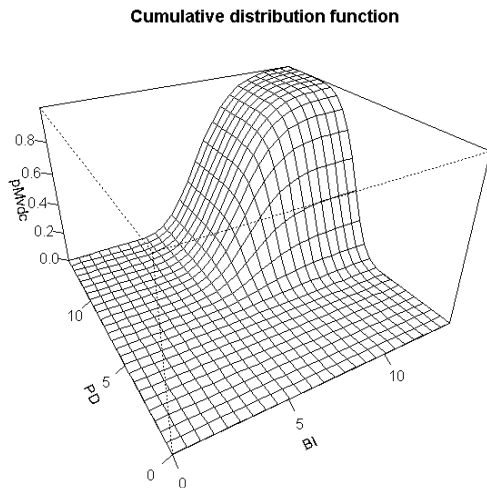
Density plot for Bivariate distribution



Density Contour plot



Στα ακόλουθα σχήματα φαίνεται η αντίστοιχη από κοινού συνάρτηση κατανομής.



Στο τελευταίο γράφημα της από κοινού συνάρτησης κατανομής (Cumulative distribution Contour plot) μπορούμε να προσδιορίσουμε που περίπου βρίσκονται οι παρατηρήσεις με τις μεγαλύτερες αποζημιώσεις, δηλαδή στην περιοχή από την προτελευταία καμπύλη και πάνω (πάνω δεξιά πλευρά του γραφήματος) βρίσκεται το 0,5% των ακραίων παρατηρήσεων.

Συνεχίζοντας, θα δημιουργήσουμε ένα προσομοιωμένο δείγμα διδιάστατων παρατηρήσεων από την διδιάστατη κατανομή που κατασκευάσαμε.

```
#τυχαίο δείγμα παρατηρήσεων από την διδιάστατη κατανομή
> sim <- rMvdc(n,dist)

#απεικόνιση του τυχαίου δείγματος και οπτική σύγκριση με τα δεδομένα
> plot(sim[,1], sim[,2], main = 'Sample dataset BI and PD', col =
'red', xlab = "", ylab = "")

> points(data, main = 'Test dataset BI and PD', col = "blue")

> legend('bottomright', c('Observed', 'Simulated'), col = c('blue',
'red'), pch=21)

> cor(data, method = "kendall")

          BI          PD
BI 1.0000000 0.04508195
PD 0.04508195 1.0000000

> cor(sim, method = "kendall")

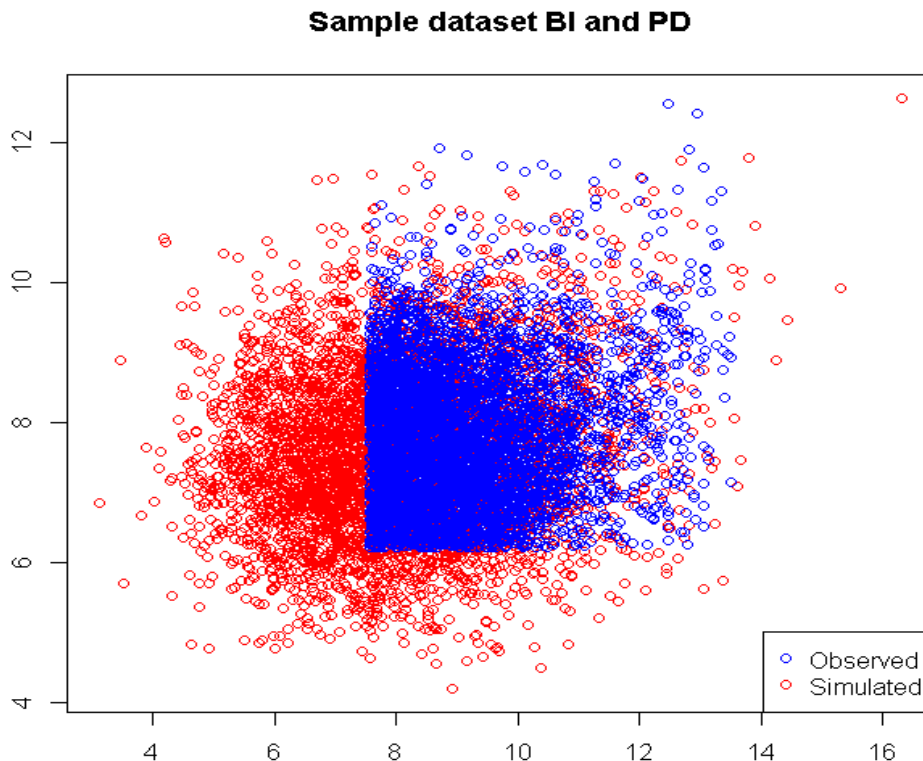
          [,1]          [,2]
```

```
[1,] 1.00000000 0.06634698
```

```
[2,] 0.06634698 1.00000000
```

Παρατηρούμε ότι ο συντελεστής συσχέτισης Kendall tau της σύζευξης είναι αρκετά κοντά στον αντίστοιχο συντελεστή που υπολογίστηκε με βάση τα προσομοιωμένα δεδομένα. Επομένως, η σύζευξη αποτύπωσε τη σχέση εξάρτησης ανεξαρτήτως των περιθωρίων συναρτήσεων.

Το προσομοιωμένο δείγμα των παρατηρήσεων (κόκκινες κουκκίδες) συγκρίνεται στο ίδιο γράφημα με τις (περικομμένες) παρατηρούμενες τιμές (μπλε κουκκίδες).



Όπως αναφέρθηκε στην αρχή του κεφαλαίου, το δείγμα το οποίο επεξεργαζόμαστε είναι περικομμένο καθώς αφαιρέθηκαν μηδενικές ζημιές ή ζημιές με μικρό κόστος. Αυτός ο περιορισμός δεν μεταφέρεται στα προσομοιωμένα δεδομένα με αποτέλεσμα να δημιουργείται η εικόνα που φαίνεται στην αριστερή πλευρά του παραπάνω γραφήματος. Ωστόσο, αν παραβλέψουμε αυτή την ασυμφωνία, παρατηρούμε ότι προσομοιωμένα δεδομένα ταιριάζουν με τα δεδομένα μας (μπλε κουκκίδες).

3.1.5 Υπολογισμός του Δεσμευμένου VaR 99.5% (Conditional VaR 99.5%)

Έχοντας πλέον, προσδιορίζει τις περιθώριες κατανομές που ταιριάζουν στα δεδομένα μας, την καταλληλότερη σύζευξη και την από κοινού συνάρτηση, θα υπολογίσουμε με βάση τα στοιχεία αυτά το VaR στο 99,5% για κάθε μία από τις περιθώριες μεταβλητές δεδομένου των τιμών της άλλης περιθώριας μεταβλητής (δεσμευμένο VaR).

```
> x=8
> qx=cCopula(c(pgamma(x, shapeBI, scaleBI), 0.995)
+ ,joeCopula(param=as.numeric(coef(fit)), dim = 2), inverse=TRUE) [2]
> qgamma(qx, shapePD, scalePD)
[1] 10.65076
```

Σύμφωνα με το παραπάνω παράδειγμα, όταν η (λογαριθμική) τιμή μιας σωματικής βλάβης BI είναι 8 ($x = 8$) (3000 ευρώ περίπου) το VaR στο 99,5% για τις υλικές ζημιές είναι 10.65, δηλαδή το αναμενόμενο ακραίο κόστος των υλικών ζημιών είναι 10.65 (γύρω στις 42200 ευρώ περίπου). Ας δούμε τώρα πως μεταβάλλεται το VaR για τις δύο μεταβλητές BI, PD για ένα σύνολο τιμών και όχι μια μεμονωμένη τιμή.

```
# VaR 99.5% δεδομένου του BI
> x=seq(1,16,0.1);qx=rep(0,length(x))
> for(i in 1:length(x))
{qx[i]=cCopula(c(pgamma(x[i], shapeBI, scaleBI), 0.995), joeCopula(param=as.numeric(coef(fit)), dim = 2), inverse=TRUE) [2]}
> plot(x, qgamma(qx, shapePD, scalePD), xlab="BI", main="PD Conditional VaR 99.5%")

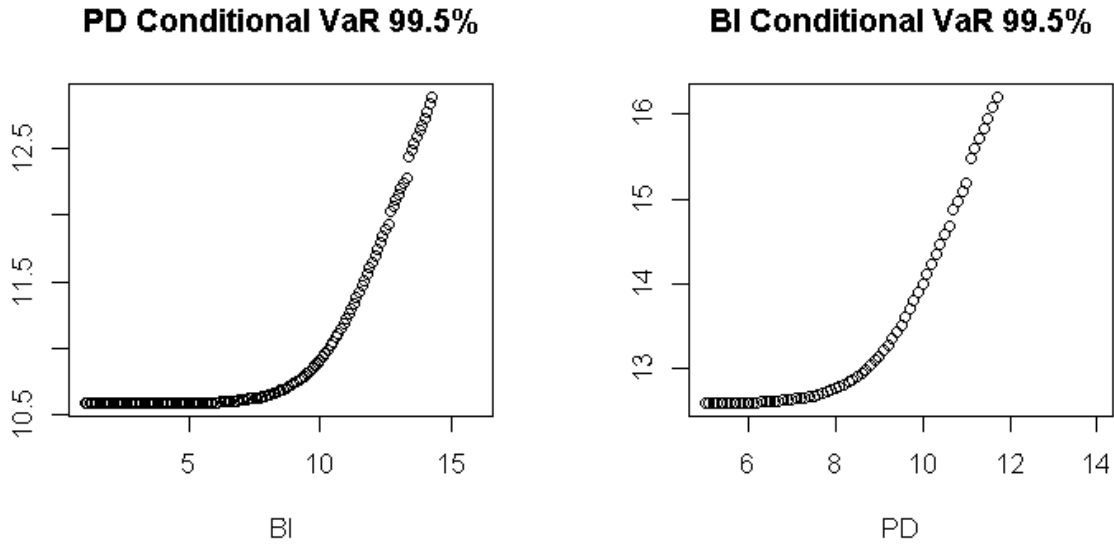
# VaR 99.5% δεδομένου του PD
> x=seq(5,14,0.1);qx=rep(0,length(x))
```

```

> for(i in 1:length(x))
{qx[i]=cCopula(c(pgamma(x[i],shapePD,scalePD),0.995),joeCopula(param=a
s.numeric(coef(fit)), dim = 2),inverse=TRUE)[2]}

> plot(x, qgamma(qx, shapeBI, scaleBI), xlab="PD", main="BI Condition-
al VaR 99.5%")

```



Μέσω του VaR ποσοτικοποιούμε τον κίνδυνο ουράς του χαρτοφυλακίου μας. Με βάση την δεσμευμένη σύζευξη Joe καταλήγουμε στις παραπάνω απεικονίσεις, όπου η συμπεριφορά του δεσμευμένου VaR για τις δύο μεταβλητές παρουσιάζει παρόμοια μορφή.

Στα παραπάνω γραφήματα π.χ. βλέπουμε ότι όταν ο λογάριθμος του κόστους των σωματικών βλαβών είναι μικρότερος του 7 περίπου ($e^7 \approx 1100\text{€}$) τότε το $\text{VaR}_{99.5\%}$ του λογαρίθμου του κόστους των υλικών ζημιών είναι περίπου 10.5 ($e^{10.5} \approx 36300\text{€}$). Όταν οι σωματικές βλάβες είναι μεγαλύτερες από 1100€ βλέπουμε ότι το $\text{VaR}_{99.5\%}$ των υλικών ζημιών αυξάνεται όπως φαίνεται στο παρακάτω πίνακα.

| BI | $BI_0 = e^{BI}$ | $\text{VaR}_{99.5\%}^{PD}$ | $\text{VaR}_{99.5\%}^{PD_0} = e^{\text{VaR}_{99.5\%}^{PD}}$ |
|----|-----------------|----------------------------|---|
| 5 | 148.41 | 10.592 | 39,797.13 |
| 6 | 403.43 | 10.597 | 40,018.21 |
| 7 | 1,096.63 | 10.613 | 40,674.64 |
| 8 | 2,980.96 | 10.651 | 42,224.67 |

| | | | |
|----|--------------|--------|------------|
| 9 | 8,103.08 | 10.739 | 46,111.61 |
| 10 | 22,026.47 | 10.916 | 55,076.04 |
| 11 | 59,874.14 | 11.220 | 74,636.13 |
| 12 | 162,754.79 | 11.648 | 114,517.16 |
| 13 | 442,413.39 | 12.153 | 189,616.71 |
| 14 | 1,202,604.28 | 12.719 | 334,064.71 |

Ανάλογα, στον παρακάτω πίνακα παρατηρούμε τις τιμές του VaR του BI ως προς τις τιμές του PD.

| PD | $PD_0 = e^{PD}$ | $VaR_{99.5\%}^{BI}$ | $VaR_{99.5\%}^{BI_0} = e^{VaR_{99.5\%}^{BI}}$ |
|------|-----------------|---------------------|---|
| 5.0 | 148.41 | 12.58 | 290,709.57 |
| 5.5 | 244.69 | 12.58 | 291,303.22 |
| 6.0 | 403.43 | 12.59 | 292,936.17 |
| 6.5 | 665.14 | 12.60 | 296,597.12 |
| 7.0 | 1,096.63 | 12.63 | 304,799.66 |
| 7.5 | 1,808.04 | 12.67 | 317,337.13 |
| 8.0 | 2,980.96 | 12.75 | 345,017.36 |
| 8.5 | 4,914.77 | 12.90 | 399,744.15 |
| 9.0 | 8,103.08 | 13.14 | 508,921.98 |
| 9.5 | 13,359.73 | 13.51 | 737,498.99 |
| 10.0 | 22,026.47 | 14.00 | 1,198,570.31 |
| 10.5 | 36,315.50 | 14.58 | 2,143,520.22 |
| 11.0 | 59,874.14 | 15.18 | 3,922,030.26 |
| 11.5 | 98,715.77 | 15.95 | 8,464,233.33 |

ΒΙΒΛΙΟΓΡΑΦΙΑ

Μπούτσικας Μιχαήλ, (2020), *Εισαγωγή στην Διαχείριση Κινδύνων*, Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης, Πανεπιστήμιο Πειραιώς.

M. Denuit, J. Dhaene, M. Goovaerts, R. Kaas, (2005), *Actuarial Theory for Dependent Risks: Measures, Orders and Models, Risks*, John Wiley & Sons, Ltd.

Henrik Hult, Filip Lindskog, (2007), *Mathematical Modeling and Statistical Methods for Risk Management*, Lecture Notes.

Alexander J. McNeil, Rüdiger Frey, Paul Embrechts, (2005), *Quantitative Risk Management: Concepts, Techniques and Tools*, Princeton University Press.

Edward W. Frees, Emiliano A. Valdez, (1998), Understanding Relationships Using Copulas, *North American Actuarial Journal*, 2(1):1-25.

W. J. Krzanowski, (1988), *Principles of Multivariate Analysis*, Oxford, Clarendon Press.

R. Dennis Cook, Mark E. Johnson, (1981), A Family of Distributions for Modelling Non-Elliptically Symmetric Multivariate Data, *Journal of the Royal Statistical Society. Series B*, Vol. 43, No. 2, 210-218.

Frank M. J., (1979), *On the Simultaneous Associativity of $F(x, y)$ and $x + y - F(x, y)$* , *Aequationes Mathematicae*, 19, 194-226.

D. G. Clayton, (1978), A Model for Association in Bivariate Life Tables and Its Application in Epidemiological Studies of Familial Tendency in Chronic Disease Incidence, *Biometrika*, Volume 65, Issue 1, 141–151, Oxford University Press.

Christian Genest, (1987), Frank's family of bivariate distributions, *Biometrika*, Volume 74, Issue 3, 549–555.

Christian Genest, Jock MacKay, (1986), The Joy of Copulas: Bivariate Distributions with Uniform Marginals, *The American Statistician*, Volume 40, No. 4, 280-283.

Joe H., (1993), Parametric families of multivariate distributions with given margins, *Journal of Multivariate Analysis* 46, 262–282.

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

Michel Denuit, Oana Purcaru, Ingrid Van Keilegom, (2004), Bivariate Archimedean copula modelling for loss-ALAE data in non-life insurance, Working Paper, STAT Discussion Paper - 0423, *ACTU Working Paper* - 2004-03.

D.A. Hennessy, H.E. Lapan, (2002), The use of Archimedean copulas to model portfolio allocations, *Mathematical Finance*, Volume 12, 143–154.

G. Kimeldorf, A. Sampson, (1975), One-parameter families of bivariate distributions with fixed marginals, *Communications in Statistics*, Volume 4, 293-301.

G. Kimeldorf, A. Sampson, (1975), Uniform representations of bivariate distributions, *Communications in Statistics*, Volume 4, 617-627.

