ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
Διατμηματικό Πρόγραμμα Μεταπτυχιακών Σπουδών
«Μαθηματικά και Εφαρμογές τους» των τμημάτων
Μαθηματικών και Εφαρμοσμένων Μαθηματικών.

Μεταπτυχιακή Εργασία

Μάρκου Μαρία

Συζεύξεις (copula) πολλών μεταβλητών

Επιβλέπων Καθηγητής : Μιχαήλ Λουλάκης

Ηράκλειο, Κρήτη, Ιούλιος 2015

Η Μεταπτυχιακή αυτή εργασία κατατέθηκε στο τμήμα Μαθηματικών της Σχολής Θετικών και Τεχνολογικών Επιστημών του Πανεπιστημίου Κρήτης , στα πλαίσια του προγράμματος μεταπτυχιακών σπουδών «Μαθηματικών και Εφαρμογές τους», στη κατεύθυνση Επιχειρησιακά Μαθηματικά, τον Ιούλιο του 2015.

Επιβλέπων Καθηγητής ήταν ο Μιχαήλ Λουλάκης.

Την επιτροπή αξιολόγησης αποτέλεσαν οι κ.κ..

Μιχαήλ Λουλάκης

Μιχαήλ Πλεξουσάκης

Γεώργιος Κοσιώρης

Ευχαριστίες

Πρώτα απ'ολα θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου , κ. Μιχάλη Λουλάκη για την πολύτιμη βοήθεια με τις γνώσεις του και την απεριόριστη υπομονή του καθ'ολη τη διάρκεια της μελέτης και συγγραφής της παρούσας μεταπτυχιακής εργασίας . Ευχαριστώ επίσης τα μέλης της επιτροπής αξιολόγησης μου , κ. Κοσιώρη και κ. Πλεξουσάκη .

Ευχαριστώ θερμά τους γονείς μου , Αναστάσιο και Αγγελική που διακριτικά βοηθούν και στηρίζουν τις προσπάθειες μου ολα αυτά τα χρόνια .

Ευχαριστώ το συμφοιτητή και κουμπάρο μου , Λευτέρη για τη στήριξη και τις ατελείωτες ώρες διαβάσματος .

Ιδιαίτερα ευχαριστώ το σύζυγο μου , Μάνο για την εμψύχωση και συμπαράσταση του αυτά τα χρόνια.

Περιεχόμενα

Κεφάλαιο 1

Συνάρτηση σύζευξης

Στο κεφάλαιο αυτό θα ορίσουμε τη συνάρτηση σύζευξης (copula), όπως ορίστηκε απο τον Sklar(1959) και θα εξετάσουμε τις βασικές ιδιότητες της. Στη συνέχεια θα δούμε κάποια παραδείγματα με διάφορες συζεύξεις καθώς και το πως συνδέονται με τη συνάρτηση κατανομής ομοιόμορφων στο [0,1] τυχαίων μεταβλητών (τ.μ.).

1.1 Βασικές έννοιες και ορισμοί

Θα ήταν χρήσιμο προτού δώσουμε τον ορισμό για την συνάρτηση σύζευξης, να ορίσουμε τις παρακάτω έννοιες. Στο εξής θα χρησιμοποιούμε το συμβολισμό $\mathbb{R}^*$, ο οποίος δεν είναι άλλος από το $[-\infty, +\infty]$.

Ορισμός 1 *Θεωρούμε δύο μη κενά κλειστά υποσύνολα A_1, A_2 του $\mathbb{R}^*$ και μια συνάρτηση $G : A_1 \times A_2 \to \mathbb{R}$. Ορίζοντας ως a_i το ελάχιστο στοιχείο του A_i για $i = 1, 2$. Η συνάρτηση G λέγεται κανονικοποιημένη (grounded) αν για κάθε $(v, z) \in A_1 \times A_2$*

$$G(a_1, z) = 0 = G(v, a_2) \tag{1.1}$$

Ορισμός 2 *Η ίδια συνάρτηση $G : A_1 \times A_2$ λέγεται 2-αύξουσα (2-increasing) αν για κάθε ορθογώνιο $[v_1, v_2] \times [z_1, z_2]$ του οποίου οι κορυφές βρίσκονται στο $A_1 \times A_2$ τ.ω. $v_1 \leq v_2, z_1 \leq z_2$*

$$V_G\big([v_1, v_2] \times [z_1, z_2]\big) = G(v_2, z_2) - G(v_1, z_2) - G(v_2, z_1) + G(v_1, z_1) \geq 0 \tag{1.2}$$

Μπορούμε τώρα να δώσουμε τον ορισμό της συνάρτησης σύζευξης C. Έστω I=[0,1].

Ορισμός 3 *Θα ονομάζουμε διδιάστατη σύζευξη (copula) μια πραγματική συνάρτηση C ορισμένη στο $I \times I$ αν έχει τις παρακάτω ιδιότητες:*
i) είναι κανονικοποιημένη, δηλαδή $C(v, 0) = C(0, z) = 0 \quad \forall (v, z) \in I \times I$
ii) είναι τέτοια ώστε $C(v, 1) = v$, $C(1, z) = z \quad \forall (v, z) \in I \times I$
iii) είναι 2-αύξουσα

Ας δούμε μερικά παραδείγματα τέτοιων συναρτήσεων.

Παράδειγμα 1
Μια συνάρτηση σύζευξης είναι η $C(v,z) = \max(v+z-1,0)$ *ορισμένη στο* $I \times I$. *Η* C *είναι σύζευξη καθώς ικανοποιεί τις ιδιότητες του ορισμού 3 διότι :*

- *είναι κανονικοποιημένη* $C(v,0) = \max(v-1,0) = 0 = \max(z-1,0) = C(0,z)$ *για κάθε* $(v,z) \in I \times I$
- *ικανοποιεί την ιδιότητα (ii) αφού* $C(v,1) = max(v,0) = v$ *και* $C(1,z) = max(z,0) = z$ *για κάθε* $(v,z) \in I \times I$
- *είναι 2-αύξουσα αφού για* $z_1 \leq z_2$ *έχουμε ότι* $max(v_2+z_1-1,0) - max(v_1+z_1-1,0) \leq max(v_2+z_2-1,0) - max(v_1+z_2-1,0)$.

Με τον ίδιο τρόπο εύκολα ελέγχει κανείς ότι η $C(v,z) = \min(v,z)$ *ορισμένη στο* $I \times I$ *ικανοποιεί τις ιδιότητες του ορισμού 3 δηλαδή είναι συνάρτηση σύζευξης.*

Ομοίως και η συνάρτηση $C(v,z) = vz$ *στο* $I \times I$ *είναι συνάρτηση σύζευξης.*

Στο εξής θα συμβολίζουμε τη συνάρτηση σύζευξης $max(v+z-1,0)$ ως C^-, την οποία θα ονομάζουμε ελάχιστη σύζευξη, την $\min(v,z)$ με C^+ και θα ονομάζουμε μέγιστη σύζευξη και τέλος τη συνάρτηση σύζευξης vz ως $C^\perp$ και θα ονομάζουμε σύζευξη γινόμενο.

Ας κάνουμε μια παρένθεση να θυμηθούμε τις ιδιότητες των συναρτήσεων κατανομής για τ.μ.. Έστω τ.μ. X , τότε η συνάρτηση κατανομής της ορίζεται ως $F_1(x) = \mathbb{P}\big[X \leq x\big]$ με πεδίο ορισμού $\mathbb{R}^*$ και έχει τις ακόλουθες ιδιότητες :

1. είναι αύξουσα συνάρτηση, δηλαδή $\forall \;\; x_1, x_2 \in \mathbb{R}^* \;\; F_1(x_1) \leq F_1(x_2)$
2. είναι δεξιά συνεχής
3. $F_1(-\infty) = 0$ και $F_1(+\infty) = 1$
4. $0 \leq F_1(x) \leq 1$ και $\mathbb{P}\big[a \leq X \leq b\big] = F_1(b) - F_1(a) \geq 0$

Έστω X , Y τ.μ. ,τότε η διδιάστατη συνάρτηση κατανομής τους ορίζεται ως $F(x,y) = \mathbb{P}\big[X \leq x, Y \leq y\big]$ με πεδίο ορισμού $\mathbb{R}^{*2}$ και έχει τις ακόλουθες ιδιότητες :

1. $F(-\infty, y) = 0 = F(x, -\infty)$ για κάθε $x, y \in \mathbb{R}^{*2}$ και $F(+\infty, +\infty) = 1$
2. είναι δεξιά συνεχής ως προς x και y και $0 \leq F(x,y) \leq 1$
3. $\forall \;\; x_1, x_2, y_1, y_2 \in \mathbb{R}^{*2}$ με $x_1 \leq x_2, y_1 \leq y_2$ ισχύει ότι

   $\mathbb{P}\big[x_1 \leq X \leq x_2, y_1 \leq Y \leq y_2\big] = F(x_2,y_2) - F(x_1,y_2) - F(x_2,y_1) - F(x_1,y_1) \geq 0$

Όπως παρατηρούμε μια διδιάστατη συνάρτηση κατανομής είναι κανονικοποιημένη εφόσον η ιδιότητα 1 ικανοποιεί τον ορισμό 1. Ενώ η ιδιότητα 3 μας εξασφαλίζει ότι μια διδιάστατη συνάρτηση κατανομής είναι 2-αύξουσα, που στην ουσία είναι η αντίστοιχη ιδιότητα της μονοδιάστατης περίπτωσης ότι η συνάρτηση κατανομής είναι αύξουσα συνάρτηση.

Γενικότερα αν μια συνάρτηση $G(x,y)$ είναι 2-αύξουσα αυτό δεν συνεπάγεται ότι είναι και μη-φθίνουσα ως προς κάθε μεταβλητη της.Όπως θα δούμε στο παρακάτω θεώρημα αυτό επιτυγχάνεται αν υποθέσουμε ότι είναι επιπλέον και κανονικοποιημένη.

Θεώρημα 1 *Μια συνάρτηση $G(v,z) : A_1 \times A_2 \to \mathbb{R}$, η οποία είναι κανονικοποιημένη και 2-αύξουσα είναι μη-φθίνουσα ως προς v και z.*
Απόδειξη:
Αρκεί να το δείξουμε για το πρώτο ορισμα v. Θα δείξουμε ότι αν $v_1 \leq v_2 \Rightarrow G(v_1,x) \leq G(v_2,x)\ \forall x \in A_2$. Η 2-αύξουσα ιδιότητα της G μας δίνει:

$$G(v_2,z_1) - G(v_1,z_1) \leq G(v_2,z_2) - G(v_1,z_2) \;\; \forall z_2 \geq z_1$$

Δηλαδή η διαφορά $G(v_2,x) - G(v_1,x)$ είναι μια μη-φθίνουσα συνάρτηση του x. Πιο συγκεκριμένα

$$G(v_2,a_1) - G(v_1,a_1) \leq G(v_2,x) - G(v_1,x) \;\; \forall x \geq a_1$$

Επειδή η G είναι κανονικοποιημένη έχουμε $G(v_1,a_1) = G(v_2,a_1) = 0$

$$G(v_1,x) \leq G(v_2,x) \;\; \forall x \in A_2 \;.\square$$

Απο τις ιδιότητες (i) , (ii) του ορισμού της σύζευξης και του θεωρήματος 1 έπεται το παρακάτω πόρισμα.

Πόρισμα 1 *Για κάθε $(v,z) \in [0,1] \times [0,1] \quad 0 \leq C(v,z) \leq 1$.*

Έχοντας ορίσει τη σύζευξη C παρατηρούμε από τον ορισμό της, ότι η από κοινού συναρτηση κατανομής δύο ομοιόμορφων τυχαίων μεταβλητών στο [0,1] ειναι μια σύζευξη. Δηλαδή αν X , Y δύο τ.μ. ομοιόμορφες στο [0,1] και F(x,y) η από κοινού συναρτηση κατανομής τους τότε η F(x,y) ειναι σύζευξη. Πράγματι αν $X \sim U_1([0,1])$ και $Y \sim U_2([0,1])$ τότε η $F(x,y) = \mathbb{P}\big[X \leq x,\ Y \leq y\big]$ ικανοποιεί τις ιδιότητες του ορισμού 3 της συνάρτησης σύζευξης διότι :

i) Η F(x,y) είναι κανονικοποιημένη
$F(x,0) = \mathbb{P}\big[X \leq x,\ Y \leq 0\big] \leq \mathbb{P}\big[Y \leq 0\big] = 0$ δηλαδή $F(x,0) = 0$ και $F(0,y) = \mathbb{P}\big[X \leq 0,\ Y \leq y\big] \leq \mathbb{P}\big[X \leq 0\big] = 0$ δηλαδή $F(y,0) = 0$

ii) Η F(x,y) ικανοποιεί την ιδιότητα (ii) του ορισμού 3
$F(x,1) = \mathbb{P}\big[X \leq x,\ Y \leq 1\big] = \mathbb{P}\big[X \leq x\big] = x$ δηλαδή $F(x,1) = x$ όμοια

$F(1,x) = \mathbb{P}\big[X \leq 1,\ Y \leq y\big] = \mathbb{P}\big[Y \leq y\big] = x$ δηλαδή $F(1,y) = y$

iii) Η F(x, y) είναι 2-αύξουσα
$F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1) = \mathbb{P}\Big[(X,Y) \in [x_1,x_2]\times[y_1,y_2]\Big] \geq 0$
όπου $[x_1,x_2] \times [y_1,y_2] \subseteq [0,1] \times [0,1]$.

Ωστόσο κάποιες συζεύξεις στις οποίες έχουμε ήδη αναφερθεί αποτελούν από κοινού κατανομή κάποιων συγκεκριμένων ομοιόμορφων τυχαίων μεταβλητών. Πιο συγκεκριμένα η σύζευξη γινόμενο $C^{\perp}$ είναι μια από κοινού συνάρτηση κατανομής δύο ανεξάρτητων ομοιόμορφων τ.μ. X,Y. Από τη θεωρία πιθανοτήτων γνωρίζουμε ότι δύο τ.μ. είναι ανεξάρτητες αν και μόνο αν η από κοινού συνάρτηση κατανομή τους είναι ίση με το γινόμενο των περιθώριων συναρτήσεων κατανομής τους. Δηλαδή $F(x,y) = \mathbb{P}\big[X \leq x,\ Y \leq y\big] = \mathbb{P}\big[X \leq x\big]\mathbb{P}\big[Y \leq y\big] = xy$. Η μέγιστη σύζευξη C^{+} είναι μια από κοινού συνάρτηση κατανομής δύο ομοιόμορφων τ.μ. X,Y όταν Y = X

Εφόσον $\mathbb{P}\big[U_1 \leq v,\ U_2 \leq z\big] = C(v,z)$ όπου $U_1 \sim U_1([0,1])$ και $U_2 \sim U_2([0,1])$ τότε οι παρακάτω πιθανότητες ομοιόμορφων τ.μ. μπορούν να γραφούν μέσω της συνάρτησης σύζευξης ως εξής:

- $\mathbb{P}\big[U_1 \leq v,\ U_2 > z\big] = v - C(v,z)$
- $\mathbb{P}\big[U_1 > v,\ U_2 \leq z\big] = z - C(v,z)$
- $\mathbb{P}\big[U_1 \leq v \mid\ U_2 \leq z\big] = \frac{C(v,z)}{z}$
- $\mathbb{P}\big[U_1 \leq v \mid\ U_2 > z\big] = \frac{v-C(v,z)}{1-z}$
- $C_{1|2} \equiv \mathbb{P}\big[U_1 \leq v \mid\ U_2 = z\big] = lim_{\Delta z \to 0^+} \frac{C(v,z+\Delta z)-C(v,z)}{\Delta z} = \frac{\partial C(v,z)}{\partial z}$
- $C_{2|1} \equiv \mathbb{P}\big[U_2 \leq z \mid\ U_1 = v\big] = \frac{\partial C(v,z)}{\partial v}$.

Όπως γνωρίζουμε από την θεωρία πιθανοτήτων αν X τ.μ. με συνάρτηση κατανομής F(x), $(X \sim F)$ τότε η αντίστροφη συνάρτηση κατανομής της F, είναι κάθε συνάρτηση F^{-1} με πεδίο ορισμού I τέτοια ώστε :

- Αν $t \in RanF$, τότε η $F^{-1}(t)$ είναι οποιοσδήποτε αριθμός $x \in \mathbb{R}^*$ έτσι ώστε $F(x) = t$, δηλαδή $\forall\ \ t \in RanF \quad F\big(F^{-1}(t)\big) = t$.
- Αν $t \notin RanF$, τότε η $F^{-1}(t)$ ορίζεται από τη γενικευμένη αντίστροφη ως : $F^{-1}(t) = inf\{x : t \leq F(x)\}\ , \forall\ t \in [0,1]$.

Τότε :

1. Αν η F είναι συνεχής τότε η τ.μ. $F(X)$ ακολουθεί την ομοιόμορφη κατανομή στο [0,1] , $F(X) \sim U([0,1])$.

2. Για κάθε U που ακολουθεί την ομοιόμορφη κατανομή στο [0,1] , ($U \sim U([0,1])$) έχουμε ότι $F^{-1}(U) \sim F$

Εφόσον η από κοινού συνάρτηση κατανομής δύο ομοιόμορφων τ.μ. στο [0,1] είναι συνάρτηση σύζευξης, τότε αυτή η σύζευξη υπολογισμένη στο σημείο $(F_1(x), F_2(y))$ είναι μια από κοινού συνάρτηση κατανομής δύο τ.μ. X , Y με περιθώριες κατανομές $F_1(x)$, $F_2(y)$ αντίστοιχα. Πράγματι αν έχουμε δύο τ.μ. X , Y με συναρτήσεις κατανομής $F_1(x)$, $F_2(y)$ αντίστοιχα και $U_1, U_2 \sim U([0,1])$ τ.ω. $F_1^{-1}(U_1) \sim F_1$ και $F_2^{-1}(U_2) \sim F_2$,οπότε :

$$\begin{aligned} C(F_1(x), F_2(y)) &= \mathbb{P}(U_1 \le F_1(x), U_2 \le F_2(y)) \\ &= \mathbb{P}(F_1^{-1}(U_1) \le x, F_2^{-1}(U_2) \le y) \\ &= \mathbb{P}(X \le x, Y \le y) \\ &= F(x, y) \end{aligned}$$

Το παραπάνω μας προϊδεάζει για την σχέση μεταξύ των συναρτήσεων κατανομής και των συζεύξεων, το οποίο θα είναι και το θέμα του θεωρήματος Sklar που θα δούμε στο κεφάλαιο 2.

1.2 Η ανισότητα *Fréchet* και η διάταξη των συζεύξεων C

Όπως αποδεικνύεται από το παρακάτω θεώρημα οι συζεύξεις είναι φραγμένες συναρτήσεις:

Θεώρημα 2 *Για κάθε συνάρτηση σύζευξης C ισχύει η παρακάτω ανισότητα :*

$$max(v+z-1,0) \le C(v,z) \le min(v,z) \quad \forall\ (v,z)\ \in [0,1]\times[0,1] \tag{1.3}$$

Απόδειξη :

- *Για την πρώτη ανισότητα $max(v+z-1,0) \le C(v,z)$ χρησιμοποιούμε την 2-αύξουσα ιδιότητα $C(v_2,z_2) - C(v_1,z_2) - C(v_2,z_1) + C(v_1,z_1) \ge 0$ η οποία ισχύει $\forall\ v_1 \le v_2, z_1 \le z_2$. Επιλέγουμε $v_2 = z_2 = 1$ και παίρνουμε $C(1,1) - C(v_1,1) - C(1,z_1) + C(v_1,z_1) \ge 0$ ή ισοδύναμα $1 - z_1 - v_1 + C(v_1,z_1) \ge 0$ που ισχύει $\forall\ v_1 \le 1, z_1 \le 1$.Συνεπώς*

  $$C(v,z) \ge v+z-1 \quad \forall\ (v,z) \in [0,1]\times[0,1]. \tag{1.4}$$

  Επιπλέον από το πόρισμα 1 έχουμε ότι

  $$C(v,z) \ge 0 \quad \forall\ \ (v,z) \in [0,1]\times[0,1]. \tag{1.5}$$

  Επομένως από τις σχέσεις 1.4 και 1.5 έπεται το ζητούμενο
  $max(v+z-1,0) \le C(v,z)$

- *Για την δεύτερη ανισότητα $C(v,z) \leq min(v,z)$ χρησιμοποιούμε το θεώρημα 1 από το οποίο έχουμε:*

$$C(v,z_1) \leq C(v,z_2) \quad \forall v \in [0,1]\ , z_1 \leq z_2$$

και

$$C(v_1,z) \leq C(v_2,z) \quad \forall z \in [0,1]\ , v_1 \leq v_2$$

Επιλέγοντας πάλι $v_2 = z_2 = 1$ έχουμε :

$$C(v,z_1) \leq v \quad \forall v \in [0,1]\ ,\ z_1 \leq 1 \tag{1.6}$$

και

$$C(v_1,z) \leq z \quad \forall z \in [0,1]\ , v_1 \leq 1 \tag{1.7}$$

Επομένως από τις σχέσεις 1.6 και 1.7 έπεται ότι $C(v,z) \leq min(v,z) \quad \forall\, (v,z) \in [0,1] \times [0,1]$. □

Το κάτω φράγμα $max(v+z-1,0)$ της ανισότητας 1.3 είναι επίσης σύζευξη, η οποία δεν είναι άλλη από την ελάχιστη σύζευξη C^- την οποία ορίσαμε στο παράδειγμα 4.2.1. Αξίζει να σημειώσουμε ότι η C^- είναι η σύζευξη που αντιστοιχεί στην από κοινού κατανομή των (X,Y) όταν $X \sim U([0,1])$ και $Y = 1-X \sim U([0,1])$. Καθώς και το άνω φράγμα $min(v,z)$ της ανισότητας 1.3 είναι η μέγιστη σύζευξη C^+. Τα όρια αυτά στο εξής θα ονομάζουμε **όρια Fréchet** και η ανισότητα 1.3 θα λέγεται **ανισότητα Fréchet** η οποία ξαναγράφεται ως :

$$C^-(v,z) \leq C(v,z) \leq C^+(v,z) \tag{1.8}$$

Η ύπαρξη ελάχιστης και μέγιστης σύζευξης μας οδηγεί να ορίσουμε μια σχέση διάταξης μεταξύ των συζεύξεων.

Ορισμός 4 *Θα λέμε οτι η σύζευξη C_1 είναι μικρότερη από την σύζευξη C_2 και θα γράφουμε $C_1 \prec C_2$ αν και μόνο αν*

$$C_1(v,z) \leq C_2(v,z) \quad \forall\ (v,z)\ \in\ [0,1] \times [0,1]$$

Ωστόσο πρέπει να ξέρουμε ότι η διάταξη όπως την ορίσαμε είναι μερική , δηλαδή δεν επιτυγχάνεται εύκολα καθώς δεν μπορούν όλες οι συζεύξεις να συγκριθούν μεταξύ τους .

Παράδειγμα 2 *Αν πάρουμε ενα κυρτό γραμμικό συνδύασμο των C^+ και C^- για παράδειγμα τον $C = \frac{1}{3}C^- + \frac{2}{3}C^+$. Τότε αυτός αποτελεί σύζευξη αφού οποιοσδήποτε κυρτός συνδυασμός συναρτήσεων κατανομής είναι συνάρτηση κατανομής.*

Τώρα την σύζευξη $C(v,z)$ που επιλέξαμε δεν γίνεται να την συγκρίνουμε με την σύζευξη γινόμενο $C^\perp$, καθώς υπάρχουν σημεία $(v,z) \in [0,1] \times [0,1]$ στα οποία $C^\perp(v,z)(v,z)$ και άλλα στα οποία ισχύει η αντίστροφη ανισότητα. Όπως για παράδειγμα στο σημείο $(\frac{1}{2},\frac{1}{2})$ ισχύει $C(\frac{1}{2},\frac{1}{2}) > C^\perp(\frac{1}{2},\frac{1}{2})$ ένω στο σημείο $(\frac{1}{4},\frac{3}{4})$ ισχύει $C(\frac{1}{4},\frac{3}{4}) < C^\perp(\frac{1}{4},\frac{3}{4})$

Κεφάλαιο 2

Το Θεώρημα του Sklar

2.1 Το Θεώρημα του Sklar και η απόδειξη του

Στο κεφάλαιο αυτό θα παρουσιάσουμε το θεώρημα Sklar και την απόδειξη του. Στη συνέχεια θα δούμε κάποια παραδείγματα και πορίσματα του θεωρήματος. Το θεώρημα Sklar είναι ένα κεντρικό θεώρημα στη θεωρία των συζεύξεων και αποτελεί βάση για πολλές, αν όχι για όλες τις χρηματοοικονομικές και στατιστικές εφαρμογές. Επιπλέον το θεώρημα αυτό αποσαφηνίζει το ρόλο που παίζουν οι συζεύξεις στη σχέση μεταξύ των από κοινού συναρτήσεων κατανομής και των περιθώριων συναρτήσεων κατανομής τους.

Θεώρημα 3 *ΘΕΩΡΗΜΑ SKLAR*

Έστω $F_1(x)$,$F_2(y)$ *περιθώριες συναρτήσεις κατανομής. Τότε για κάθε* $(x, y) \in \mathbb{R}^{*2}$ *:*

1. *Αν* C *οποιαδήποτε σύζευξη και* $F_1(x)$,$F_2(y)$ *περιθώριες συναρτήσεις κατανομής τότε η :*
$$C(F_1(x), F_2(y))$$
είναι μια από κοινού συνάρτηση κατανομής με περιθώριες συναρτήσεις $F_1(x)$,$F_2(y)$.

2. *Αντίστροφα, αν* $F(x, y)$ *είναι μια από κοινού συνάρτηση κατανομής με περιθώριες συναρτήσεις* $F_1(x)$,$F_2(y)$ *τότε υπάρχει σύζευξη* C *τέτοια ώστε για κάθε* $(x, y) \in \mathbb{R}^{*2}$ *:*
$$F(x, y) = C(F_1(x), F_2(y)) \tag{2.1}$$
Αν οι $F_1(x)$ *και* $F_2(y)$ *είναι συνεχείς συναρτήσεις τότε η σύζευξη* C *είναι μοναδική, σε αντίθετη περίπτωση η* C *ορίζεται μοναδικά στο* $RanF_1 \times RanF_2$.

(δηλαδή η C περιορισμένη στο $RanF_1 \times RanF_2$ είναι μοναδική.)

Για την απόδειξη του θεωρήματος Sklar θα χρειαστούμε τα παρακάτω λήμματα καθώς και τον ορισμό της περιθώριας συνάρτησης που ακολουθεί.

Ορισμός 5 *Οι περιθώριες συναρτήσεις της από κοινού συνάρτησης κατανομής $G : A_1 \times A_2 \to \mathbb{R}$ είναι οι συναρτήσεις $G_1 : A_1 \to \mathbb{R}$, $G_2 : A_2 \to \mathbb{R}$ και ορίζονται ως εξής :*
$G_1(x) = G(x, \bar{a_2})$ και $G_2(y) = G(\bar{a_1}, y)$ όπου $\bar{a_i}$ το μέγιστο στοιχείο του συνόλου $\bar{A_i}$.

Λήμμα 1 *Κάθε από κοινού συνάρτηση κατανομής $G : A_1 \times A_2 \to \mathbb{R}$ η οποία είναι κανονικοποιημένη και 2-αύξουσα με περιθώριες G_1, G_2 ικανοποιεί τις παρακάτω σχέσεις :*

$$|G(v_2, z) - G(v_1, z)| \leq |G_1(v_2) - G_1(v_1)| \tag{2.2}$$

για κάθε ζεύγος σημείων (v_2, z), (v_1, z) στο $A_1 \times A_2$ και

$$|G(v, z_2) - G(v, z_1)| \leq |G_2(z_2) - G_2(z_1)| \tag{2.3}$$

για κάθε ζεύγος σημείων (v, z_2), (v, z_1) στο $A_1 \times A_2$.

απόδειξη :
Περίπτωση 1: Χωρίς βλάβη υποθέτουμε ότι $v_2 > v_1$
Εφαρμόζοντας την 2-αύξουσα ιδιότητα στο ορθογώνιο $[v_1, v_2] \times [z, \bar{a_2}]$ παίρνουμε:

$$G(v_2, \bar{a_2}) - G(v_1, \bar{a_2}) - G(v_2, z) + G(v_1, z) \geq 0$$

$$\Rightarrow \quad G(v_1, z) - G(v_2, z) \geq G_1(v_1) - G_1(v_2)$$

$$G(v_2, z) - G(v_1, z) \leq G_1(v_2) - G_1(v_1)(> 0, \quad G_1 \uparrow)$$

και τα δύο μέλη της παραπάνω εξίσωσης είναι θετικά διότι η συνάρτηση G είναι μη φθίνουσα ως προς την μεταβλητή v, $\forall z \in A_2$, το ίδιο και η G_1 είναι μη φθίνουσα σαν συνάρτηση κατανομής. Επομένως :

$$|G(v_2, z) - G(v_1, z)| \leq |G_1(v_2) - G_1(v_1)|$$

Περίπτωση 2: Χωρίς βλάβη υποθέτουμε ότι $v_2 < v_1$
Ακολουθώντας την ίδια διαδικασία στο ορθογώνιο $[v_2, v_1] \times [z, \bar{a_2}]$ παίρνουμε:

$$|G(v_1, z) - G(v_2, z)| \leq |G_1(v_1) - G_1(v_2)|$$

η οποία είναι ισοδύναμη με την σχέση 2.2 που θέλαμε να δείξουμε.

Για την σχέση 2.3 παίρνοντας πάλι περιπτωσεις για το z και εφαρμόζοντας την 2-αύξουσα ιδιότητα στο ορθογώνιο $[v, \bar{a_2}] \times [z_1, z_2]$ καταλήγουμε στη ζητούμενη σχέση 2.3 . □

Λήμμα 2 *Για την συνάρτηση του λήμματος 1* $G : A_1 \times A_2 \to \mathbb{R}$ *έχουμε ότι :*

$$|G(v_1, z_1) - G(v_2, z_2)| \leq |G_1(v_1) - G_1(v_2)| + |G_2(z_1) - G_2(z_2)| \tag{2.4}$$

για κάθε $(v_1, z_1), (v_2, z_2) \in A_1 \times A_2$.

Απόδειξη :
Ξεκινώντας απο το αριστερό μέλος της ανισότητας 2.4, προσθαφαιρούμε τον όρο $G(v_2, z_1)$, *κάνουμε την τριγωνική ανισότητα*

$$\begin{aligned} |G(v_1, z_1) - G(v_2, z_2)| &= |G(v_1, z_1) - G(v_2, z_1) + G(v_2, z_1) - G(v_2, z_2)| \\ &\leq |G(v_1, z_1) - G(v_2, z_1)| + |G(v_2, z_1) - G(v_2, z_2)| \\ &\leq |G_1(v_1) - G_1(v_2)| + |G_2(z_1) - G_2(z_2)| \end{aligned}$$

όπου η τελευταία ανισότητα προκύπτει από το λήμμα 1. □

Ακολουθεί η απόδειξη του θεωρήματος Sklar.

απόδειξη :
Πρώτο μέρος του θεωρήματος. Θα δείξουμε ότι αν $RanF_1 \subseteq [0,1]$ και $RanF_2 \subseteq [0,1]$ τότε η συνάρτηση $F(x,y)$ ορισμένη από : $F(x,y) = C(F_1(x), F_2(y))$

α. είναι από κοινού συνάρτηση κατανομής

β. έχει περιθώριες συναρτήσεις $F_1(x)$, $F_2(y)$

Για τον α ισχυρισμό πρέπει να δείξουμε ότι :

1. το πεδίο ορισμού της F , $DomF = \mathbb{R}^{*2}$
2. η F είναι 2-αύξουσα
3. η F είναι κανονικοποιημένη
4. ισχύει $F(+\infty, -\infty) = 1$

Επομένως έχουμε :

- Για το (1): έχουμε ότι $DomF_1 = \mathbb{R}^*$, $DomF_2 = \mathbb{R}^*$ και αφού $RanF_1 \subseteq [0,1]$, $RanF_2 \subseteq [0,1]$ και η $F(x,y) = C(F_1(x), F_2(y))$ έπεται ότι $DomF = DomF_1 \times DomF_2 = \mathbb{R}^{*2}$.
- Για το (2): η F είναι 2-αύξουσα διότι η C είναι 2-αύξουσα
  $F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1) =$
  $C(F_1(x_2), F_2(y_2)) - C(F_1(x_1), F_2(y_2)) - C(F_1(x_2), F_2(y_1)) + C(F_1(x_1), F_2(y_1))$
  ≥ 0 για καθε ορθογώνιο $[F_1(x_1), F_1(x_2)] \times [F_2(y_1), F_2(y_2)]$
  επομένως και για κάθε ορθογώνιο $[x_1, x_2] \times [y_1, y_2]$ εφόσον οι F_1, F_2 είναι μη-φθίνουσες συναρτήσεις.

- Για το (3): η F είναι κανονικοποιημένη διότι η C είναι κανονικοποιημένη $F(-\infty, y) = C(F_1(-\infty), F_2(y)) = C(0, F_2(y)) = 0$ αντίστοιχα $F(x, -\infty) = C(F_1(x), F_2(-\infty)) = C(F_1(x), 0) = 0$.
- Για το (4): $F(+\infty, +\infty) = C(F_1(+\infty), F_2(+\infty)) = C(1, 1) = 1$ από την ιδιότητα ii) του ορισμού 3 .

Για τον β ισχυρισμό αρκεί να ελέγξουμε ότι η πρώτη περιθώρια της F , $F(x, +\infty)$, είναι πράγματι η περιθώρια συνάρτηση κατανομής $F_1(x)$. Όπου πράγματι $F(x, +\infty) = C(F_1(x), 1) = F_1(x)$. Αντίστοιχα και για την περιθώρια συνάρτηση κατανομής $F_2(y)$.

Δεύτερο μέρος του θεωρήματος.
Θεωρούμε μια από κοινού συνάρτηση κατανομής F(x,y) με περιθώριες $F_1(x)$, $F_2(y)$ και δύο σημεία $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^{*2}$. Αν υποθέσουμε ότι $F_1(x_1) = F_1(x_2)$ και $F_2(y_1) = F_2(y_2)$ τότε εφαρμόζοντας το λήμμα 2 στην F(x,y) στα σημεία $(x_1, y_1), (x_2, y_2)$ έχουμε ότι $F(x_1, y_1) = F(x_2, y_2)$. Δηλαδή για κάθε σημείο $(x, y) \in \mathbb{R}^{*2}$ η τιμή της F εξαρτάται από τις $F_1(x)$, $F_2(y)$ διαφορετικά λέμε ότι υπάρχει (μοναδικη ;;) συνάρτηση C με $DomC = RanF_1 \times RanF_2$ τ.ω. $F(x, y) = C(F_1(x), F_2(y))$. Επομένως αρκεί να δείξουμε ότι η C είναι σύζευξη. Αρκεί η $C(F_1(x), F_2(y))$ να ικανοποιεί τις ιδιότητες του ορισμού 3 :

- είναι κανονικοποιημένη καθώς
$$C(0, F_2(y)) = C(F_1(-\infty), F_2(y)) = F(-\infty, y) = 0$$
$$C(F_1(x), 0) = C(F_1(x), F_2(-\infty)) = F(x, -\infty) = 0$$
- ικανοποιεί την ιδιότητα ii) του ορισμού 3 καθώς
$$C(1, F_2(y)) = C(F_1(+\infty), F_2(y)) = F(+\infty, y) = F_2(y),$$
$$C(F_1(x), 1) = C(F_1(x), F_2(+\infty)) = F(x, +\infty) = F_1(x)$$
- είναι 2-αύξουσα διότι οι συναρτήσεις κατανομής είναι 2-αύξουσες

Άρα η C είναι σύζευξη.Αν τώρα οι $F_1(x)$, $F_2(y)$ είναι συνεχείς συναρτήσεις τότε $RanF_2 = RanF_1 = [0, 1]$ και το πεδίο ορισμού της C είναι $DomC = [0, 1] \times [0, 1]$. Αλλιώς η C ορίζεται μοναδικά στο $RanF_1 \times RanF_2$. □

2.2 Συνέπειες του θεωρήματος Sklar

Σύμφωνα με το θεώρημα του Sklar γράφοντας τη σχέση $F(x, y) = C(F_1(x), F_2(y))$ 2.1 μπορεί κάποιος να διασπάσει την από κοινού συνάρτηση κατανομής στις περιθώριες συναρτήσεις και την σύζευξη C ε.ω. η παραπάνω σχέση να παριστά τη σχέση μεταξύ της X και Y. Στις συζεύξεις η περιθώρια συμπεριφορά διαχωρίζεται σε αντίθεση με την από κοινού συνάρτηση κατανομής. Γιάυτό το λόγο οι συζεύξεις λέγονται και εξαρτημένες συναρτήσεις πολλές φορές.

Στην παράγραφο 1.2 συναντήσαμε τα **όρια Fréchet** C^- , C^+ σαν γενικά όρια για τις συζεύξεις. Η **ανισότητα Fréchet** , $C^- \leq C \leq C^+$ σαν συνέπεια του θεωρήματος Sklar γίνεται :
αν X , Y είναι τ.μ. με από κοινού συνάρτηση κατανομής F και περιθώριες κατανομές F_1, F_2, αντίστοιχα, για κάθε $x, y \in \mathbb{R}^*$

$$max(F_1(x) + F_2(y) - 1, 0) \leq F(x, y) \leq min(F_1(x), F_2(y)) \tag{2.5}$$

Η παραπάνω ανισότητα είναι γνωστή ως **Fréchet – Hoeffding ανισότητα** για από κοινού συναρτήσεις κατανομών. Επιπλέον επειδή οι C^- και C^+ είναι συζεύξεις τα παραπάνω όρια είναι κι αυτά απο κοινού συναρτήσεις κατανομών.

Για τις τ.μ. X , Y με από κοινού συνάρτηση κατανομής F και περιθώριες κατανομές F_1, F_2, αντίστοιχα, από το Θεώρημα του Sklar ισχύουν οι παρακάτω πιθανότητες :

- $\mathbb{P}(X \leq x,\ Y > y) = F_1(x) - C(F_1(x), F_2(y))$
- $\mathbb{P}(X > x,\ Y \leq y) = F_2(y) - C(F_1(x), F_2(y))$
- $\mathbb{P}(X \leq x \mid\ Y \leq y) = \frac{C(F_1(x), F_2(y))}{F_2(y)}$
- $\mathbb{P}(X \leq x \mid\ Y > y) = \frac{F_1(x) - C(F_1(x), F_2(y))}{1 - F_2(y)}$

Έπεται σαν πόρισμα του θεωρήματος Sklar ότι η σύζευξη C για την οποία $F(x, y) = C(F_1(x), F_2(y))$ μπορεί να ανασκευαστεί από τις περιθώριες F_1, F_2 και την από κοινού συνάρτηση κατανομής F χρησιμοποιώντας τη γενικευμένη αρχή της αντίστροφης. Ειδικότερα χρησιμοποιείται σαν μέθοδος για την κατασκευή της σύζευξης C.

Πόρισμα 2 *Αντίστοιχα με το 2ο μέρος του θεωρήματος Sklar, αν F(x,y) από κοινού συνάρτηση κατανομής με περιθώριες συναρτήσεις $F_1(x)$,$F_2(y)$ τότε η σύζευξη C για την οποία $F(x, y) = C(F_1(x), F_2(y))$ για κάθε $(x, y) \in \mathbb{R}^{*2}$ είναι η :*

$$C(v, z) = F(F_1^{-1}(v), F_2^{-1}(z)) \tag{2.6}$$

απόδειξη:
Θεωρούμε την σύζευξη C του 2ου μέρους του θεωρήματος Sklar για την οποία $F(x, y) = C(F_1(x), F_2(y))$. Αν ορίσουμε την αξία της $F_1(x)$ ως v και της $F_2(y)$ ως z και όταν οι συναρτήσεις κατανομής F_1 και F_2 είναι αντιστρέψιμες έχουμε :
$x = F_1^{-1}(v)$ και $y = F_2^{-1}(z)$ επομένως
$F(F_1^{-1}(v), F_2^{-1}(z)) = F(x, y) = C(F_1(x), F_2(y)) = C(v, z)$.
*Στην περίπτωση όπου η συνάρτηση κατανομής δεν είναι αντιστρέψιμη τότε ορίζουμε την **γενικευμένη αντίστροφη συνάρτηση** ως :*

$$F_i^{-1}(t) = inf\{u : F_i(u) \geq 0, 0 < t < 1\} \quad i = 1, 2 \ . \square$$

Παράδειγμα 3 *Έστω X, Y δύο τ.μ. που ακολουθούν τν εκθετική κατανομή με παραμέτρους λ_1 και λ_2 αντίστοιχα. Τότε :*
$F_1(x) = 1 - e^{-\lambda_1 x}$ *και* $F_2(y) = 1 - e^{-\lambda_2 y}$ *για* $x, y > 0$ *και* $\lambda_1, \lambda_2 > 0$
Υποθέτουμε επίσης ότι η από κοινού συνάρτηση κατανομής τους είναι η :

$$F(x,y) = max(1 - e^{-\lambda_1 x} - e^{-\lambda_2 y}, 0)$$

Αν $v = F_1(x)$ *τότε* $F_1^{-1}(v) = \frac{-\ln(1-v)}{\lambda_1}$ *, αντίστοιχα* $F_2^{-1}(z) = \frac{-\ln(1-z)}{\lambda_2}$
η σύζευξη " *για την οποία:*

$$\begin{aligned} C(F_1(x), F_2(y)) &= C(1 - e^{-\lambda_1 x}, 1 - e^{-\lambda_2 y}) \\ &= max(1 - e^{-\lambda_1 x} - e^{-\lambda_2 y}, 0) \\ &= F(x,y) \end{aligned}$$

είναι η :

$$\begin{aligned} C(v,z) &= F(F_1^{-1}(v), F_2^{-1}(z)) \\ &= F\left(\frac{-\ln(1-v)}{\lambda_1}, \frac{-\ln(1-z)}{\lambda_2}\right) \\ &= max(1 - e^{-\lambda_1 \frac{-\ln(1-v)}{\lambda_1}} - e^{-\lambda_2 \frac{-\ln(1-z)}{\lambda_2}}, 0) \\ &= max(1 - (1-v) - (1-z), 0) \\ &= max(v + z - 1, 0) \ . \end{aligned}$$

Παράδειγμα 4 *Έστω X, Y δύο τ.μ. με από κοινού συνάρτηση κατανομής την :*

$$F(x,y) = \begin{cases} \frac{(x+1)(e^y-1)}{x+2e^y-1}, & \text{αν } (x,y) \in [-1,1] \times [0,\infty] \\ 1 - e^{-y}, & \text{αν } (x,y) \in (1,\infty] \times [0,\infty] \\ 0, & \text{αλλού} \end{cases}$$

και συνεχείς περιθώριες συναρτήσεις κατανομής τις :

$$F_1(x) = \begin{cases} 0, & \text{αν } x < -1 \\ \frac{x+1}{2}, & \text{αν } x \in [-1,1] \\ 1, & \text{αν } x > 1 \end{cases} \ , \text{ και}$$

$$F_2(y) = \begin{cases} 0, & \text{αν } y < 0 \\ 1 - e^{-y}, & \text{αν } y \geq 0 \end{cases}$$

Οι αντίστροφες συναρτήσεις τους δίνονται από τις :
$F_1^{-1}(v) = 2v - 1$ *και* $F_2^{-1}(z) = -\ln(1-z)$ *για κάθε* $(v, z) \in [0, 1]$.

Η μοναδικη σύζευξη ᷉ για την οποία: $C(F_1(x), F_2(y)) = F(x, y)$ *δίνεται από την :*

$$\begin{aligned}
C(v,z) &= F(F_1^{-1}(v), F_2^{-1}(z)) \\
&= F(2v-1, -\ln(1-z)) \\
&= \frac{(2v-1+1)(e^{-\ln(1-z)}-1)}{2v-1+2e^{-\ln(1-z)}-1} \\
&= \frac{vz}{v+z-vz}
\end{aligned}$$

Καθώς όταν $0 \le v \le 1 \Rightarrow -1 \le 2v-1 \le 1$ *και* $0 \le z \le 1 \Rightarrow 0 \le -\ln(1-z) \le \infty$.

2.3 Παραδείγματα

Παράδειγμα 5 *Θεωρούμε για παράδειγμα την ελάχιστη σύζευξη* C^-, *ορισμένη στο* $RanF_1 \times RanF_2$, *τότε η*

$$C(F_1(x), F_2(y)) = max(F_1(x) + F_2(y) - 1, 0)$$

είναι από κοινού συνάρτηση κατανομής, αν οι F_1 *και* F_2 *είναι περιθώριες κατανομές, αφού :*

α) *ορίζεται για κάθε* $(x, y) \in \mathbb{R}^{*2}$

β) *είναι 2-αύξουσα αφού :*

$$\begin{aligned}
&max(F_1(x_2) + F_2(y_2) - 1, 0) - max(F_1(x_1) + F_2(y_2) - 1, 0) \\
&- max(F_1(x_2) + F_2(y_1) - 1, 0) + max(F_1(x_1) + F_2(y_1) - 1, 0) \ge 0
\end{aligned}$$

γ) *είναι κανονικοποιημένη αφού :*

$$\begin{aligned}
max(F_1(-\infty) + F_2(y) - 1, 0) = 0 \\
max(F_1(x) + F_2(-\infty) - 1, 0) = 0
\end{aligned}$$

δ) *δίνει*

$$C(F_1(+\infty), F_2(+\infty)) = max(1, 0) = 1$$

ε) *είναι απο δεξιά συνεχής, αφού είναι οι* F_1 *και* F_2 *είναι συνεχείς. Οι περιθώριες τους είναι*

$$C(F_1(x), F_2(+\infty) = max(F_1(x), 0) = F_1(x)$$

$$C(F_1(+\infty), F_2(y) = max(0, F_2(y)) = F_2(y)$$

Τα παραπάνω επαληθεύουν το 1ο μέρος του θεωρήματος Sklar . Για το 2ο μέρος του θεωρήματος Sklar ας θεωρήσουμε 2 τ.μ. X , Y ομοιόμορφες στο [0,1] με :

$$F_1(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \leq x \leq 1, \\ 1, & x > 1 \end{cases}$$

$$F_2(y) = \begin{cases} 0, & y < 0 \\ y, & 0 \leq y \leq 1, \\ 1, & y > 1 \end{cases}$$

Αν υποθέσουμε ότι η από κοινού συνάρτηση κατανομής είναι :

$$F(x,y) = \begin{cases} max(inf(x,1) + inf(y,1) - 1, 0) & , inf(x,y) > 0 \\ 0 & , \text{αλλιώς} \end{cases}$$

Αφού οι F_1 και F_2 είναι συνεχείς υπάρχει μοναδική σύζευξη C τέτοια ώστε

$$F(x,y) = C(F_1(x), F_2(y))$$

Στο παράδειγμα που ακολουθεί οι περιθώριες συναρτήσεις F_1 και F_2 δεν είναι συνεχείς.

Παράδειγμα 6 *Θεωρούμε τις ακόλουθες κατανομές :*

$$F(x,y) = \begin{cases} 0 & , x \text{ ή } y < 0 \\ \frac{1}{2}(inf(x,1) + inf(y,1)) & , 0 \leq x, y \leq 1 \end{cases}$$

$$F_1(x) = \begin{cases} 0 & , x < 0 \\ \frac{1}{2}(x+1) & , 0 \leq x \leq 1, \\ 1 & , x > 1 \end{cases}$$

$$F_2(y) = \begin{cases} 0 & , y < 0 \\ \frac{1}{2}(y+1) & , 0 \leq y \leq 1, \\ 1 & , y > 1 \end{cases}$$

Σύμφωνα με το 2ο μέρος του θεωρήματος Sklar υπάρχει σύζευξη τ.ω. $F(x,y) = C(F_1(x), F_2(y))$ για κάθε $(x,y) \in \mathbb{R}^{*2}$. Οι περιθώριες $F_1(x)$ και $F_2(y)$ δεν είναι συνεχείς, άρα η copula αυτή δεν είναι μοναδική στο $[0,1] \times [0,1]$. Το γεγονός ότι οι περιθώριες δεν είναι συνεχείς συνεπάγεται ότι $RanF_1 \times RanF_2 \neq [0,1] \times [0,1]$ καθώς $RanF_1 = \{0\} \cup [\frac{1}{2}, 1]$, αντίστοιχα $RanF_2 = \{0\} \cup [\frac{1}{2}, 1]$. Περιορίζοντας το πεδίο ορισμού στο $\{0\} \cup [\frac{1}{2}, 1] \times \{0\} \cup [\frac{1}{2}, 1]$ έχουμε ότι η σύζευξη C μπορεί να οριστεί με μοναδικό τρόπο. Έτσι η ελάχιστη σύζευξη, C^- ορισμένη στο

$\{0\} \cup [\frac{1}{2}, 1] \times \{0\} \cup [\frac{1}{2}, 1]$ *ικανοποιεί το 2ο μέρος του θεωρήματος Sklar. Πράγματι για* $\frac{1}{2} \leq x,\ y \leq 1$ *έχουμε :*

$$
\begin{aligned}
F(x,y) &= \frac{1}{2}(x+y) \\
&= max(\frac{1}{2}(x+1) + \frac{1}{2}(y+1) - 1, 0) \\
&= max(F_1(x) + F_2(y) + 1, 0) \\
&= C^{-}(F_1(x), F_2(y))
\end{aligned}
$$

Αντίστοιχα και για το σημείο (0,0) έχουμε ότι :

$$F(0,0) = 0 = C^{-}(F_1(0), F_2(0)) = max(F_1(0)+F_2(0)+1, 0) = max(\frac{1}{2}+\frac{1}{2}+1, 0).$$

Παράδειγμα 7 *Εστω δύο τ.μ. X και Y. Αν η τ.μ. X ακολουθεί την εκθετική κατανομή με παράμετρο* λ_1 *και η τ.μ. Y ακολουθεί την student κατανομή με* v *βαθμούς ελευθερίας δηλαδή* $F_1(x) = 1 - e^{-\lambda_1 x}$ *για* $x, \lambda_1 > 0$ *και*

$$F_2(y) = \int_{-\infty}^{y} \frac{\Gamma(\frac{v+1}{2})}{\sqrt{\pi v}\Gamma(\frac{v}{2})}(1+\frac{s^2}{v})^{-\frac{v+1}{2}} ds \quad \text{για } y > 0$$

όπου Γ είναι η συνάρτηση Euler. Τότε σύμφωνα με το 1ο μέρος του θεωρήματος Sklar αν χρησιμοποιήσουμε την σύζευξη $C^{\perp}$ *η από κοινού συνάρτηση κατανομής θα είναι η :*

$$
\begin{aligned}
F(x,y) &= C^{\perp}((F_1(x), F_2(y)) = F_1(x)F_2(y) \\
&= (1 - e^{-\lambda_1 x}) \int_{-\infty}^{y} \frac{\Gamma(\frac{v+1}{2})}{\sqrt{\pi v}\Gamma(\frac{v}{2})}(1+\frac{s^2}{v})^{-\frac{v+1}{2}} ds
\end{aligned}
$$

Θεωρώντας τις ίδιες περιθώριες $F_1(x), F_2(y)$ *και την ακόλουθη σύζευξη :*

$$\dot{C}(v,z) = \begin{cases} vz^{1-r} & , \text{αν } v \leq z \\ v^{1-r}z & , \text{αν } v > z \end{cases}$$

γνωστή ως Cuadras-Auge σύζευξη. Τότε με το θεώρημα του Sklar και η $\dot{C}(F_1(x), F_2(y)$ *είναι από κοινού συνάρτηση κατανομής καθώς :*

$$\dot{F}(x,y) = \dot{C}((F_1(x), F_2(y))$$

$$= \begin{cases} (1 - e^{-\lambda_1 x})(\int_{-\infty}^{y} \frac{\Gamma(\frac{v+1}{2})}{\sqrt{\pi v}\Gamma(\frac{v}{2})}(1+\frac{s^2}{v})^{-\frac{v+1}{2}} ds)^{1-r}, & \text{αν } x \leq g(y) \\ (1 - e^{-\lambda_1 x})^{1-r} \int_{-\infty}^{y} \frac{\Gamma(\frac{v+1}{2})}{\sqrt{\pi v}\Gamma(\frac{v}{2})}(1+\frac{s^2}{v})^{-\frac{v+1}{2}} ds), & \text{αν } x > g(y) \end{cases}$$

όπου το $g(y) = -\frac{1}{\lambda_1} \ln(1 - \int_{-\infty}^{y} \frac{\Gamma(\frac{v+1}{2})}{\sqrt{\pi v}\Gamma(\frac{v}{2})}(1 + \frac{s^2}{v})^{-\frac{v+1}{2}} ds)$.

Καταφέραμε λοιπόν χρησιμοποιώντας τις δύο περιθώριες συναρτήσεις κατανομής να δημιουργήσουμε δύο συναρτήσεις κατανομής για τις τ.μ. X, Y *χρησιμοποιώντας δύο διαφορετικές συζεύξεις και μπορούμε να δημιουργήσουμε τόσες όσες και οι συζεύξεις που έχουμε στη διάθεση μας. Ωστόσο οι συναρτήσεις κατανομής αυτές μας δίνουν διαφορετικές πληροφορίες για την εξάρτηση των* X *και* Y, *ειδικότερα η* F *μας δίνει την πληροφορία ότι οι τ.μ.* X *και* Y *είναι ανεξάρτητες, αφού είναι το γινόμενο των περιθώριων, ενώ η* $\dot{F}$ *όχι εκτός αν* $r = 0$.

Κεφάλαιο 3

Οι συζεύξεις σαν εξαρτημένες συναρτήσεις

Αφού οι συζεύξεις είναι εξαρτημένες συναρτήσεις, μας επιτρέπουν να ορίσουμε την ανεξαρτησία και την τέλεια εξάρτηση με εναν απλό τρόπο.
Στο εξής θα θεωρούμε ότι οι ($F_1(x)$ και $F_2(y)$) είναι συνεχείς συναρτήσεις ούτως ώστε οι συζεύξεις να ορίζονται μοναδικά στο [0,1].

3.1 Ανεξάρτητες τ.μ.

Όπως ξέρουμε δύο συνεχείς τ.μ. X και Y είναι ανεξάρτητες αν και μόνο αν η από κοινού συνάρτηση κατανομής τους ισούται με το γινόμενο των περιθώριων κατανομων, δηλαδή αν $F(x,y) = F_1(x)F_2(y)$.Από αυτό και το θεώρημα του Sklar αποδεικνύεται το παρακάτω πόρισμα.

Πόρισμα 3 *Οι συνεχείς τ.μ. X και Y είναι ανεξάρτητες αν και μόνο αν έχουν την σύζευξη γινόμενο* $C^{\perp}$ *.*
απόδειξη: *η απόδειξη είναι προφανής αφού αν οι τ.μ.X, Y έχουν την σύζευξη* $C^{\perp}$ *τότε* $C^{\perp}(F_1(x), F_2(y)) = F_1(x)F_2(y) \Leftrightarrow F(x,y) = F_1(x)F_2(y)$ *από Sklar.*$\square$

3.2 Η ευστάθεια των συζεύξεων σε μονότονους μετασχηματισμούς των τ.μ. X και Y

Μια χρήσιμη ιδιότητα των συζεύξεων είναι ότι για αυστηρά μονότονους μετασχηματισμούς των τ.μ. X και Y, οι συζεύξεις είτε παραμένουν αμετάβλητες είτε η αλλαγή τους γίνεται με προβλέψιμο τρόπο. Όπως ξέρουμε αν η συνάρτηση κατανομής μιας τ.μ. X είναι συνεχής και αν $a : \mathbb{R}^* \to \mathbb{R}$ είναι μια αυστηρά μονότονη συνάρτηση της οποίας το πεδίο ορισμού περιέχει το $RanX$, τότε η συνάρτηση κατανομής της τ.μ. $a(X)$ είναι επίσης συνεχής . Ας δούμε πρώτα την περίπτωση όπου οι συναρτήσεις a_1, a_2 των τ.μ. X,Y είναι γνησίως αύξουσες συναρτήσεις :

Θεώρημα 4 *Έστω X και Y συνεχείς τ.μ. με περιθώριες συναρτήσεις κατανομής F_1, F_2 και σύζευξη C. Αν a_1, a_2 γνησίως αύξουσες συναρτήσεις στο $RanX$ και $RanY$ αντίστοιχα , τότε οι τ.μ. $a_1(X), a_2(Y)$ με περιθώριες συναρτήσεις κατανομής $H_1(x) = F_1(a_1^{-1}(x))$, $H_2(y) = F_2(a_2^{-1}(y))$ και από κοινού συνάρτηση κατανομής :*

$$H(u,t) = \mathbb{P}(a_1(X) \leq u, a_2(Y) \leq t)$$

έχουν την ίδια σύζευξη C επομένως $H(u,t) = C(H_1(u), H_2(t))$
(δηλαδή $C_{a_1(X),a_2(Y)} = C_{X,Y}$.)

Απόδειξη :
Για τις τ.μ. $a_1(X), a_2(Y)$ με αντίστοιχες περιθώριες

$$\begin{aligned} H_1(x) &= \mathbb{P}(a_1(X) \leq x) = \mathbb{P}(X \leq a_1^{-1}(x)) = F_1(a_1^{-1}(x)) \\ H_2(y) &= \mathbb{P}(a_2(Y) \leq y) = \mathbb{P}(Y \leq a_2^{-1}(y)) = F_2(a_2^{-1}(y)) \end{aligned}$$

και από κοινού κατανομή $H(u,t) = \mathbb{P}(a_1(X) \leq u \ , \ a_2(Y) \leq t)$.
Ορίζουμε ως $\tilde{C}$ την σύζευξη των τ.μ. $a_1(X), \ a_2(Y)$ και θα δείξουμε ότι $\tilde{C} = C$. Επομένως για κάθε $(u,t) \in \mathbb{R}^{*2}$ έχουμε ότι :

$$\begin{aligned} H(u,t) &= \tilde{C}(H_1(u), H_2(t)) \\ &= \mathbb{P}(a_1(X) \leq u, a_2(Y) \leq t) \\ &= \mathbb{P}(X \leq a_1^{-1}(u), Y \leq a_2^{-1}(t)) \\ &= C(F_1(a_1^{-1}(u)), F_2(a_2^{-1}(t))) \\ &= C(H_1(u), H_2(t)) \end{aligned}$$

κι επειδή οι τ.μ. X και Y είναι συνεχείς , $RanH_1 = RanH_2 = [0,1]$ προκύπτει ότι $\tilde{C} = C$ στο $[0,1] \times [0,1]$. □

Ενώ όταν τουλάχιστον μια από τις συναρτήσεις a_1, a_2 είναι γνησίως φθίνουσα, έχουμε σαν αποτελέσματα η σύζευξη των τ.μ. $a_1(X)$ και $a_2(Y)$ να είναι ένας απλός μετασχηματισμός της σύζευξης των τ.μ. X και Y. Συγκεκριμένα έχουμε :

Θεώρημα 5 *Έστω X και Y συνεχείς τ.μ. με περιθώριες συναρτήσεις κατανομής F_1, F_2 και σύζευξη C. Αν a_1, a_2 γνησίως μονότονες συναρτήσεις στο $RanX$ και $RanY$ αντίστοιχα. Τότε για τις τ.μ. $a_1(X), a_2(Y)$ με περιθώριες συναρτήσεις κατανομής H_1 , H_2 και από κοινού συνάρτηση κατανομής :*
$H(u,t) = \mathbb{P}(a_1(X) \leq u \ , \ a_2(Y) \leq t)$ έχουμε ότι :

1. *Αν* a_1 *γνησίως αύξουσα και* a_2 *γνησίως φθίνουσα, τότε :*
$$H(u,t) = H_1(u) - C(H_1(u), 1 - H_2(t))$$
2. *Αν* a_1 *γνησίως φθίνουσα και* a_2 *γνησίως αύξουσα, τότε :*
$$H(u,t) = H_2(t) - C(1 - H_1(u), H_2(t))$$
3. *Αν* a_1 *και* a_2 *γνησίως φθίνουσες, τότε :*
$$H(u,t) = H_1(u) + H_2(t) - 1 + C(1 - H_1(u), 1 - H_2(t))$$

Απόδειξη για το 1ο :
Για τις τ.μ. $a_1(X), a_2(Y)$ *με αντίστοιχες περιθώριες*
$$H_1(x) = \mathbb{P}(a_1(X) \leq x) = \mathbb{P}(X \leq a_1^{-1}(x)) = F_1(a_1^{-1}(x))$$
$$H_2(y) = \mathbb{P}(a_2(Y) \leq y) = \mathbb{P}(Y \geq a_2^{-1}(y)) = 1 - \mathbb{P}(Y \leq a_2^{-1}(y) = 1 - F_2(a_2^{-1}(y))$$
και από κοινού κατανομή $H(u,t) = \mathbb{P}(a_1(X) \leq u, a_2(Y) \leq t)$.
Ορίζουμε ως $\tilde{C}$ *την σύζευξη των τ.μ.* $a_1(X),\ a_2(Y)$ *και θα δείξουμε ότι* $\tilde{C} = C$.
Επομένως για κάθε $(u,t) \in \mathbb{R}^{*2}$ *έχουμε ότι :*
$$\begin{aligned}
H(u,t) &= \tilde{C}(H_1(u), H_2(t)) \\
&= \mathbb{P}(a_1(X) \leq u\ ,\ a_2(Y) \leq t) \\
&= \mathbb{P}(X \leq a_1^{-1}(u)\ ,\ Y \geq a_2^{-1}(t)) \\
&= \mathbb{P}(X \leq a_1^{-1}(u)) - \mathbb{P}(X \leq a_1^{-1}(u), Y \leq a_2^{-1}(t)) \\
&= F_1(a_1^{-1}(u)) - C(F_1(a_1^{-1}(u)), F_2(a_2^{-1}(t))) \\
&= H_1(u) - C(H_1(u), 1 - H_2(t))
\end{aligned}$$

Παρόμοια βγαίνουν τα (2) και (3). □

Αξίζει να σημειώσουμε ότι σε κάθε μια από τις παραπάνω περιπτώσεις ο τύπος της σύζευξης είναι ανεξάρτητος από την επιλογή των συναρτήσεων a_1, a_2.

3.3 Πυκνότητα σύζευξης

Κάθε σύζευξη σαν από κοινου συνάρτηση κατανομής χαρακτηρίζεται από μια συνάρτηση πυκνότητας, η οποία ορίζεται ως εξής:

Ορισμός 6 *Η συνάρτηση πυκνότητας* $c(v,z)$ *η οποία συνδέεται με την σύζευξη* $C(v,z)$ *ορίζουμε να είναι η:*
$$c(v,z) = \frac{\partial^2 C(v,z)}{\partial v \partial z}$$
όπου (για όποια (v,z) *) αυτή η μικτή παράγωγος υπάρχει .*

Θεώρημα 6 *Η πυκνότητα $c(v,z)$ υπάρχει σχεδόν παντού στο εσωτερικό του $[0,1]^2$ και είναι μη αρνητική.*
(για την απόδειξη του παραπάνω θεωρήματος δείτε [CLV])

Για συνεχείς τ.μ. X,Y η πυκνότητα της σύζευξης C συνδέεται με την από κοινού συνάρτηση πυκνότητας και τις περιθώριες συναρτήσεις πυκνότητας με τον ακόλουθη σχέση βάσει του θεωρήματος Sklar :

$$c(F_1(x), F_2(y)) = \frac{f(x,y)}{f_1(x)f_2(y)} \tag{3.1}$$

Η παραπάνω σχέση προκύπτει εύκολα καθώς από το θεώρημα του Sklar και την πυκνότητα c της σύζευξης έχουμε :

$$\begin{aligned} c(F_1(x), F_2(y)) &= \frac{\partial^2 C(F_1(x), F_2(y))}{\partial F_1(x) \partial F_2(y))} \\ &= \frac{\partial^2 F(x,y)}{\partial F_1(x) \partial F_2(y))} \\ &= \frac{\frac{\partial^2 F(x,y)}{\partial x \partial y}}{\frac{\partial F_1(x) \partial F_2(y))}{\partial x \partial y}} \\ &= \frac{f(x,y)}{f_1(x)f_2(y)} \ . \end{aligned}$$

Για κάθε σύζευξη C έστω

$$C(v,z) = A_C(v,z) + S_C(v,z) \quad \text{όπου}$$

$$A_C(v,z) = \int_0^v \int_0^z \frac{\partial^2 C(u,t)}{\partial u \partial t} dudt = \int_0^v \int_0^z c(u,t) dudt$$

$$S_C(v,z) = C(v,z) - A_C(v,z)$$

Όταν $C \equiv A_C$ στο $[0,1]^2$ γεγονός που ισχύει, αφού αν θεωρήσουμε σαν από κοινού συνάρτηση κατανομής τη σύζευξη C τότε έχει από κοινού πυκνότητα $\frac{\partial^2 C(v,z)}{\partial v \partial z}$, τότε η C θα λέγεται **απολύτως συνεχής**. Ενώ όταν $C \equiv S_C$ στο $[0,1]^2$ που ισχύει όταν η από κοινού πυκνότητα είναι 0 σχεδόν παντου στο $[0,1]^2$, $c(v,z) = 0$, τότε η C θα λέγεται **ιδιόμορφη**(singular). Αλλιώς, όταν η σύζευξη C δεν είναι ούτε απολύτως συνεχής ούτε ιδιόμορφη θα λέμε ότι έχει μια απολύτως συνεχή συνιστώσα A_C και μια ιδιόμορφη συνιστώσα S_C. Σε αυτή την περίπτωση ούτε η A_C ούτε η S_C είναι σύζευξη, επειδή καμία δεν έχει σαν περιθώρια την ομοιόμορφη στο [0,1]. Επιπλέον το C-μέτρο της απολύτως συνεχούς συνιστώσας είναι το $A_C(1,1)$ και της ιδιόμορφης συνιστώσας είναι το $S_C(1,1)$.

Όπως ο φορέας μιας από κοινού συνάρτησης κατανομής H είναι το συμπλήρωμα της ένωσης όλων των ανοικτών υποσυνόλων του $\mathbb{R}^2$(με H- μέτρο μηδέν), ο φορέας μιας σύζευξης είναι το συμπλήρωμα της ένωσης όλων των ανοικτών υποσυνόλων του $[0,1]^2$ (με C-μέτρο μηδέν). Όταν ο φορέας της C είναι το $[0,1]^2$ θα λέμε ότι η C έχει πλήρη φορέα. Ωστόσο πολλές συζεύξεις που έχουν πλήρη φορέα έχουν και απολύτως συνεχή συνιστώσα και ιδιόμορφη .

Παράδειγμα 8

- *Η γινόμενο σύζευξη* $C^{\perp}$ *είναι απολύτως συνεχής, διότι για κάθε* $(v,z) \in [0,1]^2$ *έχουμε :*

$$\begin{aligned} A_{C^{\perp}}(v,z) &= \int_0^v \int_0^z \frac{\partial^2 C^{\perp}(u,t)}{\partial u \partial t} dudt \\ &= \int_0^v \int_0^z \frac{\partial^2 ut}{\partial u \partial t} dudt = \int_0^v \int_0^z 1 dudt \\ &= vz = C^{\perp}(v,z) \end{aligned}$$

- *Για την ελάχιστη σύζευξη* C^- *προκύπτει ότι είναι ιδιόμορφη διότι:*

$$c(v,z) = \frac{\partial^2 C(v,z)}{\partial v \partial z} = \frac{\partial^2 max(v+z-1,0)}{\partial v \partial z} = 0$$

- *Για την μέγιστη σύζευξη* C^+ *προκύπτει επίσης ότι είναι ιδιόμορφη διότι:*

$$c(v,z) = \frac{\partial^2 C(v,z)}{\partial v \partial z} = \frac{\partial^2 min(v,z)}{\partial v \partial z} = 0$$

3.4 Συμμονοτονία-αντιμονοτονία

Μέχρι τώρα έχουμε δει πως όταν δύο τ.μ. X και Y έχουν από κοινού συνάρτηση κατανομής ίση με την γινόμενο σύζευξη, δηλαδή όταν $F(x,y) = C^{\perp}(F_1(x), F_2(y))$ τότε οι X και Y είναι ανεξάρτητες. Ένα ενδιαφέρον ερώτημα είναι για το τι συμβαίνει όταν δύο τ.μ. έχουν από κοινού συνάρτηση κατανομής ίση με την μέγιστη ή την ελάχιστη σύζευξη, όταν δηλαδή $F(x,y) = C^+(F_1(x), F_2(y))$ ή $F(x,y) = C^-(F_1(x), F_2(y))$. Για να απαντήσουμε στο ερώτημα αυτό θα χρειαστεί να εισάγουμε μια νέα έννοια, που είναι αυτή της συμμονοτονίας (και αντιμονοτονίας). Η οποία όπως θα δούμε έχει να κάνει με την εξάρτηση των τ.μ. X και Y. Πιο συγκεκριμένα θα δούμε πως δύο τ.μ. θα είναι συμμονότονες δηλαδή θα έχουν τέλεια θετική εξάρτηση αν και μόνο αν έχουν την μέγιστη σύζευξη C^+. Αντίστοιχα δύο τ.μ. θα είναι αντιμονότονες δηλαδή θα έχουν τέλεια αρνητική εξάρτηση αν και μόνο αν έχουν την ελάχιστη σύζευξη C^-. Για να τεκμηριώσουμε τα παραπάνω θα δώσουμε τους παρακάτω ορισμούς.

Ορισμός 7 *Το σύνολο $A \subset \mathbb{R}^{*2}$ λέγεται συμμονότονο αν και μόνο αν για κάποιο* $(x_1, y_1), (x_2, y_2) \in A$ *,είτε* $\begin{cases} x_1 \leq x_2 \\ y_1 \leq y_2 \end{cases}$ *είτε* $\begin{cases} x_1 \geq x_2 \\ y_1 \geq y_2 \end{cases}$

Έτσι αν ένα σύνολο A είναι συμμονότονο και $(x_1, y_1), (x_2, y_2) \in A$, τότε αν $x_1 < x_2$ (ή $x_2 < x_1$) τότε θα πρέπει οπωσδήποτε να ισχύει $y_1 \leq y_2$ (ή $y_2 \leq y_1$). Είναι προφανές πως κάθε υποσύνολο συμμονότονου είναι επίσης συμμονότονο σύνολο.

Παράδειγμα 9 *Έστω το σύνολο $A = \{(x, y) \mid y = 2x \in \mathbb{R}\}$ τότε το σύνολο $A \subseteq \mathbb{R}^{*2}$ είναι συμμονότονο αφού για οποιαδήποτε $x_1, x_2 \in A$ με*

$$x_1 \leq x_2 \quad \Rightarrow 2x_1 \leq 2x_2 \Rightarrow y_1 \leq y_2$$

Ορισμός 8 *Ενα τυχαίο διάνυσμα (X,Y) θα λέγεται συμμονότονο ή τέλεια θετικά εξαρτημένο αν και μόνο αν υπάρχει συμμονότονο σύνολο $A \subset \mathbb{R}^{*2}$ τ.ω.*

$$\mathbb{P}((X, Y) \in A) = 1$$

Η έννοια του συμμονότονου τυχαίου διανύσματος (X,Y) καθώς και η έννοια της θετικής εξάρτησης έχει να κάνει με την σχέση των τιμών που παίρνει η τ.μ X και τις τιμές που παίρνει η τ.μ.Y. Αν η μία παίρνει μεγάλες τίμες τότε παιρνει και η άλλη ομοίως μεγάλες τιμες.
Το θεώρημα που ακολουθεί μας δίνει περισσότερες εναλλακτικές για τον χαρακτηρισμό ενός ενός τυχαίου διανύσματος ως συμμονότονου.

Θεώρημα 7 *Ένα τυχαίο διάνυσμα (X,Y) με (συνεχείς) περιθώριες κατανομές F_1, F_2 και από κοινού συνάρτηση κατανομής F είναι συμμονότονο αν και μόνο αν ισχύει ένα από τα παρακάτω :*

(1) *Το (X,Y) έχει συμμονότονο φορέα*

(2) *Για κάθε $(x, y) \in \mathbb{R}^{*2}$ ισχύει $F(x, y) = min(F_1(x), F_2(y))$*

(3) $C(v, z) = C^+(v, z)$

(4) $(X, Y) \sim (F_1^{-1}(U), F_2^{-1}(U))$ *όπου* $U \sim U([0, 1])$

(5) $(X, Y) \sim (F_1^{-1}(F_2(Y)), F_2^{-1}(F_1(X)))$

απόδειξη:
$(1) \rightarrow (2)$
Υποθέτουμε ότι το (X,Y) έχει συμμονότονο φορέα A και έστω τα σύνολα

$$A_1 \equiv \{(s, t) \in A : s \leq x\} \quad \textit{και} \quad A_2 \equiv \{(s, t) \in A : t \leq y\}$$

τότε $\mathbb{P}\big(X \leq x, Y \leq y\big) = \mathbb{P}\big((X, Y) \in A_1 \cap A_2\big)$ *έτσι λόγω συμμονοτονίας του* A

- *είτε* $A_1 \subset A_2 \Rightarrow A_1 \cap A_2 = A_1$ *και*

$$F(x,y) = \mathbb{P}\big((X,Y) \in A_1 = F_1(x)\big)$$

- *είτε* $A_1 \supset A_2 \Rightarrow A_1 \cap A_2 = A_2$ *και*

$$F(x,y) = \mathbb{P}\big((X,Y) \in A_2 = F_2(y)\big)$$

επομένως έχουμε ότι $F(x,y) = \min\big(F_1(x), F_2(y)\big)$.
$(2) \to (3)$
Έχουμε ότι για κάθε $(x,y) \in \mathbb{R}^{*2}$ *ισχύει* $F(x,y) = min(F_1(x), F_2(y))$, *δηλαδή* $F(x,y) = C^+(F_1(x), F_2(y))$. *Από το θεώρημα του Sklar όμως έχουμε ότι αν* $F(x,y)$ *από κοινού συνάρτηση κατανομής με περιθώριες* F_1, F_2 *τότε υπάρχει μοναδική σύζευξη* C *τ.ω.* $F(x,y) = C(F_1(x), F_2(y))$ *έπομένως αν υποθέσουμε ότι οι* F_1, F_2 *είναι συνεχείς τότε πρέπει να ισχύει* $C(v,z) = C^+(v,z)$.
$(3) \to (4)$
Θέλουμε να δείξουμε ότι $F(x,y) = C(F_1(x), F_2(y)) = min(F_1(x), F_2(y))$
συνεπάγεται ότι:
$F(x,y) = \mathbb{P}(F_1^{-1}(U) \le x, F_2^{-1}(U) \le y)$ *όπου* $U \sim U([0,1])$
όμως

$$\begin{aligned}\mathbb{P}(F_1^{-1}(U) \le x, F_2^{-1}(U) \le y) &= \mathbb{P}(U \le F_1(x), U \le F_2(y)) \\ &= \mathbb{P}(U \le min(F_1(x), F_2(y)) \\ &= min(F_1(x), F_2(y))\end{aligned}$$

αφού η τ.μ. $U \sim U([0,1])$.
$(4) \to (5)$
Έχουμε ότι $(X,Y) \sim (F_1^{-1}(U), F_2^{-1}(U))$ *με* $U \sim U([0,1])$.
Άρα αφού $X = F_1^{-1}(U)$ *και* $F_1(x)$ *συνεχής* $U = F_1(X)$
και $Y = F_2^{-1}(U) = F_2^{-1}(F_1(X))$ *συνεπάγεται ότι* $Y = F_2^{-1}(F_1(X))$.

Όμοια για το Y , $Y = F_2^{-1}(U)$ *και* $F_2(x)$ *συνεχής* $U = F_2(Y)$
και $X = F_1^{-1}(U) = F_1^{-1}(F_2(Y))$ *συνεπάγεται ότι* $X = F_1^{-1}(F_2(Y))$.
$(5) \to (1)$
Δεδομένου ότι $(X,Y) \sim (F_1^{-1}(F_1(X)), F_2^{-1}(F_2(Y)))$
έχουμε ότι οι πιθανές τιμές που θα πάρει το τυχαίο διάνυσμα (X, Y) θα είναι :

$$S = \{(x, F_2^{-1}(F_1(x))) : x \in \mathbb{R}^*\}$$

και αφού οι (F_1, F_2) *είναι περιθώριες συναρτήσεις κατανομής, επομένως είναι μη φθίνουσες συναρτήσεις τότε το παραπάνω σύνολο* S *είναι συμμονότονο άρα και το ζεύγος (X, Y) είναι συμμονότονο.* □

Από το προηγούμενο θεώρημα είδαμε πως οι τ.μ. X και Y έχουν την μέγιστη σύζευξη C^+ αν και μόνο αν το διάνυσμα (X,Y) έχει συμμονότονο φορέα. Το να έχει συμμονότονο φορέα το διάνυσμα (X,Y) όμως είναι ισοδύναμο με το να πούμε ότι το σύνολο $S = \{(x, F_2^{-1}(F_1(x))) : x \in \mathbb{R}^*\}$ είναι συμμονότονο, δηλαδη είναι μη φθίνον ως προς κάθε συνιστώσα. Δηλαδή αν $x_1, x_2 \in \mathbb{R}^*$ και $x_1 < x_2$ τότε $F_2^{-1}(F_1(x_1)) \leq F_2^{-1}(F_1(x_2))$. Επομένως μπορούμε να πούμε ότι η τ.μ. Y είναι αύξουσα συνάρτηση της τ.μ. X αν και μόνο αν έχουν την μέγιστη σύζευξη C^+. Ενώ αν οι τ.μ. X και Y είναι ομοιόμορφες τ.μ. στο $[0,1]$ και έχουν τη μέγιστη σύζευξη C^+ τότε $\mathbb{P}(X = Y) = 1$.

Ο αντίστοιχος ορισμός για τα αντιμονότονα σύνολα και τυχαία διανύσματα είναι ο ακόλουθος.

Ορισμός 9 *Το σύνολο $A \subset \mathbb{R}^{*2}$ λέγεται αντιμονότονο αν και μόνο αν για κάποιο* $(x_1, y_1), (x_2, y_2) \in A$ *,είτε* $\begin{cases} x_1 \leq x_2 \\ y_1 \geq y_2 \end{cases}$ *είτε* $\begin{cases} x_1 \geq x_2 \\ y_1 \leq y_2 \end{cases}$

Έτσι αν ένα σύνολο A είναι αντιμονότονο και $(x_1, y_1), (x_2, y_2) \in A$, τότε αν $x_1 < x_2$ (ή $x_2 < x_1$) τότε θα πρέπει οπωσδήποτε να ισχύει $y_1 \geq y_2$ (ή $y_2 \geq y_1$) . Είναι προφανές πως κάθε υποσύνολο αντιμονότονου είναι επίσης αντιμονότονο σύνολο.

Παράδειγμα 10 *Έστω το σύνολο $B = \{(x, y) \mid y = -2x \in \mathbb{R}\}$ τότε το σύνολο $B \subseteq \mathbb{R}^{*2}$ είναι αντιμονότονο αφού για οποιαδήποτε $x_1, x_2 \in B$ με*

$$x_1 \leq x_2 \quad \Rightarrow -2x_1 \geq -2x_2 \Rightarrow y_1 \geq y_2$$

Ορισμός 10 *Ενα τυχαίο διάνυσμα (X, Y) θα λέγεται αντιμονότονο ή τέλεια αρνητικά εξαρτημένο αν και μόνο αν υπάρχει αντιμονότονο σύνολο $A \subset \mathbb{R}^{*2}$ τ.ω.*

$$\mathbb{P}((X, Y) \in A) = 1$$

Υπάρχει το αντίστοιχο θεώρημα για αντιμονότονα τυχαία διανύσματα (X,Y), το οποίο είναι:

Θεώρημα 8 *Ένα τυχαίο διάνυσμα (X, Y) με (συνεχείς) περιθώριες κατανομές F_1, F_2 και από κοινού συνάρτηση κατανομής F είναι αντιμονότονο αν και μόνο αν ισχύει ένα από τα παρακάτω :*

(1) *Το (X, Y) έχει αντιμονότονο φορέα*

(2) *Για κάθε $(x, y) \in \mathbb{R}^{*2}$ ισχύει $F(x, y) = max(F_1(x), F_2(y))$*

(3) $C(v, z) = C^-(v, z)$

(4) $(X, Y) \sim (F_1^{-1}(U), 1 - F_2^{-1}(U))$ *όπου $U \sim U([0, 1])$*

(5) $(X, Y) \sim (F_1^{-1}(1 - F_2(Y)), F_2^{-1}(1 - F_1(X)))$

Η απόδειξη του παραπάνω θεωρήματος είναι παρόμοια με αυτη του θεωρήματος 7.

Αντίστοιχα οι τ.μ. X και Y έχουν την ελάχιστη σύζευξη C^- αν και μόνο αν το διάνυσμα (X,Y) έχει αντιμονότονο φορέα. Δηλαδή το σύνολο $S = \{(x, F_2^{-1}(F_1(x))) : x \in \mathbb{R}^*\}$ είναι αντιμονότονο, άρα είναι μη αύξον ως προς κάθε συνιστώσα. Δηλαδή αν $x_1, x_2 \in \mathbb{R}^*$ και $x_1 < x_2$ τότε $F_2^{-1}(F_1(x_1)) \geq F_2^{-1}(F_1(x_2))$ και το αντίστροφο. Επομένως μπορούμε να πούμε ότι η τ.μ. Y είναι φθίνουσα συνάρτηση της τ.μ. X αν και μόνο αν έχουν την ελάχιστη σύζευξη C^-. Ενώ αν οι τ.μ. X και Y είναι ομοιόμορφες τ.μ. στο $[0, 1]$ και έχουν τη ελάχιστη σύζευξη C^- τότε $\mathbb{P}(X + Y = 1) = 1$.

3.5 Σύζευξη επιβίωσης (survival copula)

Σε πολλές χρηματοοικονομικές εφαρμογές οι τ.μ. που μας ενδιαφέρουν συμβολίζουν τη διάρκεια ζωής ή επιβίωσης κάποιων φυσικών προσώπων ή αντικειμένων σε κάποιο πληθυσμο. Μπορεί για παράδειγμα μια τ.μ. να συμβολίζει τον χρόνο πτώχευσης ή επιβίωσης μια εταιρείας. Η πιθανότητα επιβίωσης πέρα απο το χρόνο x για μια τ.μ. X, με περιθώρια συνάρτηση κατανομής $F_1(x)$ θα δίνεται από : $\bar{F}_1(x) = \mathbb{P}[X > x] = 1 - F_1(x)$ και θα λέγεται περιθώρια συνάρτηση επιβίωσης (marginal survival function). Για ένα τυχαίο διάνυσμα (X,Y) με από κοινού συνάρτηση κατανομής F, η από κοινού συνάρτηση επιβίωσης (joint survival function), $\bar{H}$, δίνεται από: $\bar{H}(x, y) = \mathbb{P}(X > x, Y > y)$. Ενώ οι περιθώριες της $\bar{F}$ είναι οι $\bar{F}_1$, $\bar{F}_2$.
Το ερώτημα που τίθεται είναι αν υπάρχει σχέση που να συνδέει την από κοινού συνάρτηση επιβίωσης $\bar{F}$ με τις περιθώριες συναρτήσεις επιβίωσης $\bar{F}_1$, $\bar{F}_2$ ανάλογη με αυτή που προκύπτει απο το θεώρημα του Sklar, την $F(x, y) = C(F_1(x), F_2(y))$. Για να απαντήσουμε σε αυτό το ερώτημα υποθέτουμε ότι η σύζευξη των τ.μ. X και Y είναι η C. Τότε :

$$\begin{aligned}\bar{F}(x, y) &= 1 - \bar{F}_1(x) - \bar{F}_2(y) + F(x, y) \\ &= \bar{F}_1(x) + \bar{F}_2(y) - 1 + C(F_1(x), F_2(y)) \\ &= \bar{F}_1(x) + \bar{F}_2(y) - 1 + C(1 - \bar{F}_1(x), 1 - \bar{F}_2(y))\end{aligned}$$

Δηλαδή

$$\bar{F}(x, y) = \bar{F}_1(x) + \bar{F}_2(y) - 1 + C(1 - \bar{F}_1(x), 1 - \bar{F}_2(y)) \tag{3.2}$$

Μπορούμε τώρα να δώσουμε τον ορισμό της σύζευξης επιβίωσης (survival copula), η οποία ορίζεται να είναι :

Ορισμός 11 *Η σύζευξη επιβίωσης (survival copula), $\bar{C}$ συνδέεται με την σύζευξη C και δίνεται από :*

$$\bar{C}(v, z) = v + z - 1 + C(1 - v, 1 - z) \tag{3.3}$$

Παρατηρήσεις

1. Η σύζευξη επιβίωσης $\bar{C}$ είναι σύζευξη καθώς ικανοποιεί όλες τις ιδιότητες του ορισμού της σύζευξης C.

   - είναι κανονικοποιημένη διότι : $\bar{C}(v,0) = v - 1 + C(1-v,1) = v - 1 + 1 - v = 0$
     όμοια για την z ισχύει $\bar{C}(0,z) = 0$.
   - είναι τέτοια ώστε : $\bar{C}(v,1) = v + 1 - 1 + C(1-v,0) = v + 0 = v$
     όμοια για την z ισχύει $\bar{C}(1,z) = z$.
   - είναι 2-αύξουσα διότι :

     $$\begin{aligned}&\bar{C}(v_2,z_2) - \bar{C}(v_1,z_2) - \bar{C}(v_2,z_1) + \bar{C}(v_1,z_1)\\ &= v_2 + z_2 - 1 + C(1-v_2,1-z_2) - v_1 - z_2 + 1 - C(1-v_1,1-z_2)\\ &- v_2 - z_1 + 1 - \bar{C}(1-v_2,1-z_1) + v_1 + z_1 - 1 + C(1-v_1,1-z_1)\\ &= C(1-v_2,1-z_2) - C(1-v_1,1-z_2) - C(1-v_2,1-z_1) + C(1-v_1,1-z_1) \geq 0\end{aligned}$$

     διότι η C είναι 2-αύξουσα σε κάθε ορθογώνιο που βρίσκεται εξ'ολοκλήρου στο $[0,1]^2$, άρα και στο ορθογώνιο $[1-v_2, 1-v_1] \times [1-z_2, 1-z_1] \quad (\subseteq [0,1]^2)$ όταν $[v_1,v_2] \times [z_1,z_2] \subseteq [0,1]^2$.

2. Αφού η σύζευξη επιβίωσης $\bar{C}$ είναι σύζευξη ικανοποιεί την ανισότητα *Fréchet*, δηλαδή $C^-(v,z) \leq \bar{C}(v,z) \leq C^+(v,z)$. Επιπλέον η ελάχιστη σύζευξη επιβίωσης ,$\bar{C}^-$, συμπίπτει με την ελάχιστη σύζευξη C^-, το ίδιο ισχύει για τη γινόμενο και τη μέγιστη συζευξη, δηλαδή :

   $$\bar{C}^-(v,z) = C^-(v,z) \quad , \quad \bar{C}^\perp(v,z) = C^\perp(v,z), \quad \bar{C}^+(v,z) = C^+(v,z)$$

3. Αν U_1, U_2 δύο ομοιόμορφες τ.μ. στο [0,1] με σύζευξη C, τότε η σύζευξη επιβίωσης $\bar{C}$ υπολογισμένη στο $(1-v, 1-z)$ μας δίνει την πιθανότητα οι U_1, U_2 να είναι μεγαλύτερες από v, z αντίστοιχα, δηλαδή ότι $\bar{C}(1-v,1-z) = \mathbb{P}(U_1 > v, U_2 > z)$. Πράγματι ισχύει διότι :

   $$\begin{aligned}\bar{C}(1-v,1-z) &= 1 - v + 1 - z - 1 + C(v,z)\\ &= 1 - \mathbb{P}(U_1 \leq v) + 1 - \mathbb{P}(U_2 \leq z) - 1 + \mathbb{P}(U_1 \leq v, U_2 \leq z)\\ &= \mathbb{P}(U_1 > v) + \mathbb{P}(U_2 > z) - 1 + \mathbb{P}(U_1 \leq v, U_2 \leq z)\\ &= 1 + \mathbb{P}(U_1 > v, U_2 > z) - 1\\ &= \mathbb{P}(U_1 > v, U_2 > z).\end{aligned}$$

Είναι σύνηθες να διαχωρίζεται η σύζευξη επιβίωσης από την από κοινού συνάρτηση επιβίωσης για ομοιόμορφες τ.μ. για το λόγο αυτό παραθέτουμε τον επόμενο ορισμό :

Ορισμός 12 *Η από κοινού συνάρτηση επιβίωσης για (τυπικά) ομοιόμορφες τ.μ. U_1, U_2 με σύζευξη C , συμβολίζεται ως $\not{C}$ και ορίζεται από την παρακάτω πιθανότητα :*

$$\not{C}(v,z) = \mathbb{P}(U_1 > v, U_2 > z)$$

Έπεται από τον ορισμό ότι :

$$C = \bar{C}(1-v, 1-z) = 1 - v - z + C(v, z)$$

Επανερχόμαστε στο ερώτημα τέθηκε προηγουμένως. Παρατηρούμε ότι η σύζευξη επιβίωσης ,$\bar{C}$ υπολογισμένη στο $(\bar{F}_1(x), \bar{F}_2(y))$ μας δίνει το δεύτερο μέλος της εξίσωσης 3.2, αφού $\bar{C}(\bar{F}_1(x), \bar{F}_2(y)) = \bar{F}_1(x)+\bar{F}_2(y)-1+C(1-\bar{F}_1(x), 1-\bar{F}_2(y))$. Επομένως μπορούμε να πούμε ότι υπάρχει μια σχέση που συνδέει την από κοινού συνάρτηση επιβίωσης $\bar{F}$ με τις περιθώριες συναρτήσεις επιβίωσης $\bar{F}_1$, $\bar{F}_2$ και είναι η εξής : $\bar{F}(x, y) = \bar{C}(\bar{F}_1(x), \bar{F}_2(y))$.

Επομένως το θεώρημα του Sklar μπορεί να αναδιατυπωθεί σε όρους σύζευξης επιβίωσης $\bar{C}$, από κοινού συνάρτηση επιβίωσης $\bar{F}$ και των περιθώριων συναρτήσεων επιβίωσης $\bar{F}_1$, $\bar{F}_2$ ως εξής:

Θεώρημα 9 *Έστω $\bar{F}_1(x)$,$\bar{F}_2(y)$ περιθώριες συναρτήσεις επιβίωσης. Τότε για κάθε $(x, y) \in \mathbb{R}^{*2}$:*

1. *Αν $\bar{C}$ οποιαδήποτε σύζευξη επιβίωσης και $\bar{F}_1(x)$,$\bar{F}_2(y)$ περιθώριες συναρτήσεις επιβίωσης τότε η :*
$$\bar{C}(\bar{F}_1(x), \bar{F}_2(y))$$
είναι μια από κοινού συνάρτηση επιβίωσης με περιθώριες συναρτήσεις $\bar{F}_1(x)$,$\bar{F}_2(y)$.

2. *Αντίστροφα, αν $\bar{F}(x, y)$ είναι μια από κοινού συνάρτηση επιβίωσης με περιθώριες συναρτήσεις $\bar{F}_1(x)$,$\bar{F}_2(y)$ τότε υπάρχει σύζευξη $\bar{C}$ τέτοια ώστε για κάθε $(x, y) \in \mathbb{R}^{*2}$:*
$$\bar{F}(x, y) = \bar{C}(\bar{F}_1(x), \bar{F}_2(y)) \tag{3.4}$$
Αν οι $F_1(x)$ και $F_2(y)$ είναι συνεχείς συναρτήσεις τότε η σύζευξη επιβίωσης $\bar{C}$ είναι μοναδική, σε αντίθετη περίπτωση η $\bar{C}$ ορίζεται μοναδικά στο $Ran\bar{F}_1 \times Ran\bar{F}_2$.

Η απόδειξη του παραπάνω θεωρήματος είναι ανάλογη του θεωρήματος Sklar.

Μια σημαντική παρατήρηση που μπορούμε να κάνουμε όσον αφορά στο παραπάνω θεώρημα είναι η εξής: στο τρίτο μέρος του θεωρήματος 5 η από κοινού συνάρτηση κατανομής των τ.μ. $a_1(X)$, $a_2(Y)$ όταν οι συναρτήσεις a_1, a_2 είναι γνησίως φθίνουσες εκφράζεται μέσω της σύζευξης επιβίωσης $\bar{C}$. Πιο συγκεκριμένα είχαμε δει για την από κοινού συνάρτηση κατανομής των τ.μ. $a_1(X)$, $a_2(Y)$, $H(u, t) = H_1(u) + H_2(t) - 1 + C(1 - H_1(u), 1 - H_2(t)) = \bar{C}(H_1(u), H_2(t))$.

Κεφάλαιο 4

Μέτρα συμφωνίας (measures of concordance)

Γενικότερα γνωρίζουμε ότι δύο τ.μ. είναι εξαρτημένες όταν δεν είναι ανεξάρτητες. Ενώ η έννοια της ανεξαρτησίας είναι ¨ξεκάθαρη' η έννοια της εξάρτησης όχι . Ωστόσο υπάρχουν πολλοί τρόποι εξάρτησης όπως η γραμμική συσχέτιση , η εξάρτηση από τα άκρα και η έννοια της συμφωνίας (concordance). Στο κεφάλαιο αυτό θα ορίσουμε τα μέτρα συμφωνίας τα οποία σχετίζονται με αυτά από τα οποία τα κυριότερα είναι οι συντελεστές Kendall's τ και Spearman's ρ_S. Καθώς η συνάρτηση σύζευξης συνδέει τη διδιάστατη συνάρτηση κατανομής με τις περιθώριες η σύζευξη θα περικλείει συγκεκριμένες πτυχές της σχέσης εξάρτησης μεταξύ δύο τ.μ. . Επιπλέον τα μέτρα αυτά είναι χρήσιμα στην μη παραμετρική στατιστική γιατί δεν κάνουν κάποιες υποθέσεις ως προς την κατανομή του διανύσματος (X, Y) . Επομένως η έννοια της εξάρτησης και τα μέτρα συμφωνίας είναι ιδιότητα των συζεύξεων. Θα τελειώσουμε με την παρουσίαση κάποιων οικογενειών συναρτήσεων σύζευξης.

4.1 Βασικές έννοιες και ορισμοί

Η έννοια της συμφωνίας (ή αρμονίας) με την πιο χαλαρή έννοια έχει να κάνει με το γεγονός ότι η πιθανότητα οι τ.μ. X και Y να παίρνουν ΄ μεγάλες ΄ τιμές (ή μικρές) είναι μεγάλη ενώ η πιθανότητα η τ.μ. X να παίρνει μεγάλες τιμές και η τ.μ. Y να παίρνει μικρές τιμές ή αντίστροφα είναι μικρή. Γενικότερα θα λέμε ότι οι τ.μ. X και Y είναι σύμφωνες (concordant) όταν μεγάλες τιμές της τ.μ. X συνδέονται με μεγάλες τιμές της τ.μ. Y και μικρές τιμές της τ.μ. X συνδέονται με μικρές τιμές της τ.μ. Y .
Επίσημα ως μέτρο συμφωνίας μεταξύ των τ.μ. X και Y με σύζευξη C θα συμβολίζουμαι ως $M_{X,Y}$ ή M_C και χαρακτηρίζουμε από τον παρακάτω ορισμό.

Ορισμός 13 *Το μέτρο $M_{X,Y} = M_C$ είναι μέτρο συμφωνίας μεταξύ των τ.μ. X και Y με σύζευξη C αν και μόνο αν ισχύουν οι παρακάτω ιδιότητες :*

1. *Το $M_{X,Y}$ ορίζεται για ζεύγος τ.μ. X και Y (πληρότητα) και είναι πραγματικός αριθμός*
2. *Είναι κανονικοποιημένο μέτρο , δηλαδή $-1 \leq M_{X,Y} \leq 1$*
3. *Είναι συμμετρικό , δηλαδή $M_{X,Y} = M_{Y,X}$*
4. *Αν οι τ.μ. X και Y είναι ανεξάρτητες , τότε $M_{X,Y} = 0$*
5. $M_{-X,Y} = M_{X,-Y} = -M_{X,Y}$
6. *Συγκλίνει όταν η σύζευξη C συγκλίνει (κατά σημείο)*
   Δηλαδή αν $\{(X_n, Y_n)\}$ είναι μια ακολουθία τ.μ. με σύζευξη C_n , και

   $$\lim_{n\to+\infty} C_n(v,z) = C(v,z) \qquad \forall (v,z) \in I \times I$$

   τότε

   $$\lim_{n\to+\infty} M_{Xn,Yn} = M_{X,Y}$$

7. *Ακολουθεί την διάταξη των συζεύξεων , δηλαδή*
   αν $C_1 \prec C_2$ τότε $M_{C_1} \leq M_{C_2}$

Από τον παραπάνω ορισμό συμπεραίνουμε ότι θα υπάρχει μια σταθερότητα σε γνησίως αύξουσων μετασχηματισμών καθώς και ότι θα υπάρχει μια αντστοιχία στα φράγματα του μέτρου M σε σχέση με την συμμονοτονία και την αντιμονοτονία. Πράγμα το οποίο μας επιβεβαιώνουν τα παρακάτω θεωρήματα.

Θεώρημα 10 *Αν a_1, a_2 είναι αύξουσες συναρτήσεις στο $RanF_1, RanF_2$ τότε για το μέτρο συμφωνίας των τ.μ. $a_1(X), a_2(Y)$ θα ισχύει*
$M_{a1(X),a2(Y)} = M_{X,Y}$.
Απόδειξη :
Για τις τ.μ. $a_1(X), a_2(Y)$ βάσει του θεωρήματος 4 έχουμε ότι έχουν κοινή σύζευξη δηλαδή $C_{a_1(X),a_2(Y)} = C_{X,Y}$. Ισοδύναμα έχουμε ότι $C_{a_1(X),a_2(Y)} \leq C_{X,Y}(1)$ και $C_{a_1(X),a_2(Y)} \geq C_{X,Y}(2)$. Τότε από την ιδιότητα 7 του ορισμού 13 οι σχέσεις (1), (2) μας δίνουν για τα μέτρα συμφωνίας ότι $M_{a1(X),a2(Y)} \leq M_{X,Y}$ και $M_{a1(X),a2(Y)} \geq M_{X,Y}$. Επομένως $M_{a1(X),a2(Y)} = M_{X,Y}$ □

Παρόμοια αποδεικνύεται ότι αν a_1, a_2 φθίνουσες συναρτήσεις στο $RanF_1, RanF_2$ για τα μέτρα συμφωνίας των τ.μ. $a_1(X), a_2(Y)$ ισχύει και πάλι ότι $M_{a1(X),a2(Y)} = M_{X,Y}$.
Ωστόσο η συμμονοτονία και η αντιμονοτονία συνδέει τα φράγματα για το μέτρο συμφωνίας M με την μέγιστη σύζευξη C^+ και την ελάχιστη σύζευξη C^- .

Θεώρημα 11 *Αν οι τ.μ. X, Y είναι συμμονότονες τότε $M_{X,Y} = 1$ ενώ αν οι τ.μ. X, Y είναι αντιμονότονες τότε $M_{X,Y} = -1$*
(για την απόδειξη του παραπάνω θεωρήματος δείτε [Sc])

Ενδιαφέρον παρουσιάζει το παρακάτω θεώρημα :

Θεώρημα 12 *Έστω φραγμένη , ασθενώς μονότονη , περιττή συνάρτηση f με πεδίο ορισμού $[-\frac{1}{2}, \frac{1}{2}]$ τότε το*

$$k \int\int_{I^2} f(v - \frac{1}{2}) f(z - \frac{1}{2}) dC(v, z) \tag{4.1}$$

με $k^{-1} = \int_I f^2(u - \frac{1}{2}) du$ είναι ένα μέτρο συμφωνίας .

Καθορίζοντας την συνάρτηση f προκύπτουν κάποια πολύ γνωστά μέτρα συμφωνίας. Για παράδειγμα αν ορίσουμε σαν $f(u) = u$ η σχέση 4.1 του θεωρήματος 12 μας δίνει τον συντελεστή Spearman's ρ_s ένα μέτρο συμφωνίας το οποίο θα ορίσουμε στη συνέχεια. Ενώ αν ορίσουμε σαν $f(u) = sgn(u)$ προκύπτει ο συντελεστής Blomqvist's β ο οποίος ορίζεται ως

$$q = 4C(\frac{1}{2}, \frac{1}{2}) - 1 \tag{4.2}$$

Ωστόσο υπάρχουν κάποια πολύ γνωστά μέτρα συμφωνίας τα οποία δεν προκύπτουν από τη σχέση 4.1 του θεωρήματος 12 όπως ο συντελεστής Kendall's τ τον οποίο θα ορίσουμε στη συνέχεια και ο συντελεστής Gini's γ ο οποίος ορίζεται ως

$$\gamma = 2 \int\int_{I^2} (|v + z - 1| - |v - z|) dC(v, z) \tag{4.3}$$

4.2 Δημοφιλή μέτρα συμφωνίας

4.2.1 Ο συντελεστής Kendall's τ

Ο συντελεστής Kendall's τ ως μέτρο συμφωνίας διατυπώθηκε πρώτη φορά σύμφωνα με τον Nelsen (1991) από τον Fechner γύρω στο 1900 και επαναδιατυπώθηκε από τον Kendall το 1938.

Ο συντελεστής Kendall's τ δίνεται από τον παρακάτω ορισμό :

Ορισμός 14 *Έστω οι τ.μ. X, Y με σύζευξη C ο συντελεστής Kendall's τ ή τ_C ορίζεται ως :*

$$\tau = 4 \int\int_{I^2} C(v, z) dC(v, z) - 1 \tag{4.4}$$

Αποδεικνύεται ότι ο Kendall's τ 'μετρά' τη διαφορά μεταξύ πιθανότητας της συμφωνίας (concordance) και ασυμφωνίας (discordance) δύο ανεξάρτητων τυχαίων διανυσμάτων (X_1, Y_1) και (X_2, Y_2) με ίδια από κοινού συνάρτηση κατανομής F και σύζευξη C. Στο εξής θα λέμε ότι τα τυχαία διανύσματα (X_1, Y_1) και (X_2, Y_2) είναι 'σύμφωνα' concordant αν $X_1 > X_2$ όταν $Y_1 > Y_2$ είτε $X_1 < X_2$ όταν $Y_1 < Y_2$ και άσύμφωνα' discordant στην αντίθετη περίπτωση. Έτσι έχουμε το παρακάτω θεώρημα.

Θεώρημα 13 *Έστω (X_1, Y_1) και (X_2, Y_2) διανύσματα ανεξάρτητων και όμοια κατανεμημένων τ.μ. με σύζευξη C τότε*

$$\tau = \tau_C = \mathbb{P}\big[(X_1 - X_2)(Y_1 - Y_2) > 0\big] - \mathbb{P}\big[(X_1 - X_2)(Y_1 - Y_2) < 0\big] \tag{4.5}$$

Απόδειξη :
Οι τ.μ. είναι συνεχείς με κοινή σύζευξη C έτσι έχουμε ότι

$$\begin{aligned}
\mathbb{P}\big[(X_1 - X_2)(Y_1 - Y_2) < 0\big] &= 1 - \mathbb{P}\big[(X_1 - X_2)(Y_1 - Y_2) > 0\big] \\
\mathbb{P}\big[(X_1 - X_2)(Y_1 - Y_2) > 0\big] &= \mathbb{P}\big[X_1 < X_2, Y_1 < Y_2\big] + \mathbb{P}\big[X_1 > X_2, Y_1 > Y_2\big] \\
&= 2\int\int_{\mathbb{I}^2} C(v,z)dC(v,z)
\end{aligned}$$

διότι

$$\begin{aligned}
\mathbb{P}\big[X_1 > X_2, Y_1 > Y_2\big] &= \mathbb{P}\big[X_2 < X_1, Y_2 < Y_1\big] \\
&= \int\int_{\mathbb{R}^2} \mathbb{P}\big[X_2 \le x, Y_2 \le y\big] dC(F_1(x), F_2(y)) \\
&= \int\int_{\mathbb{R}^2} C\big(F_1(x), F_2(y)\big) dC(F_1(x), F_2(y)) \\
&= \int\int_{\mathbb{I}^2} C\big(v,z\big)dC(v,z)
\end{aligned}$$

Ομοίως , εφόσον οι τ.μ. έχουν κοινή σύζευξη C :

$$\mathbb{P}\big[X_1 < X_2, Y_1 < Y_2\big] = \int\int_{\mathbb{I}^2} C\big(v,z\big)dC(v,z)$$

Επομένως χρησιμοποιώντας τις παραπάνω σχέσεις έχουμε :

$$\begin{aligned}
\tau = \tau_C &= \mathbb{P}\big[(X_1 - X_2)(Y_1 - Y_2) > 0\big] - \mathbb{P}\big[(X_1 - X_2)(Y_1 - Y_2) < 0\big] \\
&= 2\mathbb{P}\big[(X_1 - X_2)(Y_1 - Y_2) > 0\big] - 1 \\
&= 4\int\int_{\mathbb{I}^2} C\big(v,z\big)dC(v,z) - 1. \quad \square
\end{aligned}$$

Αποδεικνύεται ότι ο συντελεστής Kendall's τ είναι ένα μέτρο συμφωνίας καθώς ικανοποιεί όλες τις ιδιότητες του ορισμού 13. Εστιάζοντας περισσότερο στην δεύτερη ιδιότητα δηλαδή ότι για τον Kendall's τ ισχύει $-1 \leq \tau \leq 1$ συνδυάζοντας ωστόσο και το θεώρημα 11 έχουμε ότι το κάτω φράγμα του Kendall's τ επιτυγχάνεται για αντιμονότονες τ.μ. ενώ το άνω φράγμα του για συμμονότονες τ.μ. . Έτσι συνδυάζοντας και τα θεωρήματα 7 και 8 του τρίτου κεφαλαίου προκύπτει το παρακάτω συμπέρασμα :

$$\tau = 1 \Leftrightarrow C = C^{-} \tag{4.6}$$

$$\tau = -1 \Leftrightarrow C = C^{+} \tag{4.7}$$

Όπου $C = C^{+}$ η μέγιστη σύζευξη που είδαμε στο παράδειγμα του πρώτου κεφαλαίου και αντίστοιχα $C = C^{-}$ η ελάχιστη σύζευξη.
Μια ακόμα σημαντική παρατήρηση που θα μπορούσαμε να κάνουμε για τον συντελεστή Kendall's τ αν θεωρήσουμε τις ομοιόμορφες τ.μ. U_1, U_2 με από κοινού συνάρτηση κατανομής C τότε το διπλό ολοκλήρωμα του ορισμού 4.4 δεν είναι τίποτα άλλο από την αναμενόμενη τιμή της συνάτησης $C(U_1, U_2)$. Επομένως βάσει της παραπάνω παρατήρησης ο συντελεστής Kendall's τ ισοδυναμεί με :

$$\tau = 4\mathbb{E}[C(U_1, U_2)] - 1$$

και

$$-1 \leq 4\mathbb{E}[C(U_1, U_2)] - 1 \leq 1$$

Δηλαδή ο συντελεστής Kendall's τ είναι μια κανονικοποιημένη αναμενόμενη τιμή.

Ενώ όταν η σύζευξη C είναι απολύτως συνεχής δηλαδή όταν $C(v, z) = A_C(v, z) = \int_0^v \int_0^z \frac{\partial^2 C(u,t)}{\partial u \partial t} dudt$, το διαφορικό

$$dC = \frac{\partial^2 C(v, z)}{\partial v \partial z} dvdz$$

μπορεί να αντικατασταθεί στον τύπο 4.4 για τον υπολογισμό του συντελεστή Kendall's τ. Ενώ όταν η σύζευξη C είναι ιδιόμορφη, δηλαδή $A_C(v, z) = 0$ ή όταν έχει μια απολύτως συνεχή A_C και μια ιδιόμορφη συνιστώσα S_C για τον υπολογισμό του συντελεστή Kendall's τ μπορεί να χρησιμοπιηθεί το παρακάτω θεώρημα.

Θεώρημα 14 *Ο συντελεστής Kendall's τ μπορεί να υπολογιστεί ως :*

$$\tau = 1 - 4 \int\int_{I^2} \frac{\partial C(v, z)}{\partial v} \frac{\partial C(v, z)}{\partial z} dvdz \tag{4.8}$$

Η απόδειξη του παραπάνω θεωρήματος έπεται από το παρακάτω λήμμα :

Λήμμα 3 *Έστω σύζευξη C τότε*

$$\int\int_{I^2} C(v,z)dC(v,z) + \int\int_{I^2} \frac{\partial C(v,z)}{\partial v}\frac{\partial C(v,z)}{\partial z}dvdz = \frac{1}{2}$$

Παράδειγμα 11 *Ας θυμηθούμε τη σύζευξη γινόμενο $C^{\perp}$,τη σύζευξη ανεξάρτητων τ.μ. με $C^{\perp}(v,z) = vz$ όπως έχουμε δει στο παράδειγμα 8 είναι απολύτως συνεχής και*

$$\frac{\partial^2 C^{\perp}(v,z)}{\partial v \partial z} = \frac{\partial^2 vz}{\partial v \partial z} = 1$$

επομένως

$$dC^{\perp} = \frac{\partial^2 C^{\perp}(v,z)}{\partial v \partial z}dvdz = dvdz$$

έτσι ο συντελεστής Kendall's τ για τη σύζευξη γινόμενο $C^{\perp}$ υπολογίζεται ως :

$$\begin{aligned}
\tau_{C^{\perp}} &= 4\int_0^1\int_0^1 C^{\perp}(v,z)dC^{\perp}(v,z) - 1 \\
&= 4\int_0^1\int_0^1 vzdvdz - 1 \\
&= 4\big(\int_0^1 z(\frac{1}{2} - 0)dz\big) - 1 \\
&= 4\frac{1}{2}(\frac{1}{2} - 0) - 1 \\
&= 1 - 1 = 0
\end{aligned}$$

Πριν δούμε το παρακάτω θεώρημα ας θυμηθούμε τη σύζευξη επιβίωσης (survival copula) , $\bar{C}$ όπως είδαμε στον ορισμό 11 η οποία συνδέεται με την σύζευξη C με την παρακάτω σχέση :

$$\bar{C}(v,z) = v + z - 1 + C(1-v, 1-z) \tag{4.9}$$

Τότε ,

Θεώρημα 15 *Ο συντελεστής Kendall's τ της σύζευξης C με της σύζευξης επιβίωσης $\bar{C}$ συμπίπτουν :*

$$\tau_C = \tau_{\bar{C}}$$

Αν ωστόσο κάποιος θέλει να υπολογίσει έναν εκτιμητή για τον συντελεστή Kendall's τ έχοντας ένα τυχαίο δείγμα n μεταβλητών

$$(X_i, Y_i) \quad \text{για} \quad i = 1, ..., n$$

ορίζοντας τις δείκτριες μεταβλητές όπως ο Gibbons το 1992

$$A_{ij} \equiv sgn(X_i - X_j)(Y_i - Y_j)$$

Τότε υπολογίζοντας τη μέση τιμή της μεταβλητής A_{ij} βάσει του θεωρήματος 13 έχουμε τον συντελεστή Kendall's τ αφού

$$\mathbb{E}(A_{ij}) = (+1)\mathbb{P}\big[(X_i - X_j)(Y_i - Y_j) > 0\big] + (-1)\mathbb{P}\big[(X_i - X_j)(Y_i - Y_j) < 0\big] = \tau$$

Έτσι ένας αμερόληπτος εκτιμητής του συντελεστή Kendall's τ είναι ο λεγόμενος δειγματικός Kendall's τ :

$$\hat{\tau} = \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{j>i} A_{ij} \tag{4.10}$$

Ο εκτιμητής μπορεί να κατασκευαστεί να είναι και συνεπής.

4.2.2 Ο συντελεστής Spearman's ρ_S

Ειδαμε πρώτη φορά στην αρχή του κεφαλαίου μετά την παρουσίαση του θεωρήματος 12 πως μπορεί να προκύψει ο συντελεστής Spearman's ρ_S . Θα ορίσουμε τον συντελεστή Spearman's ρ_S ως ένα δεύτερο μέτρο συμφωνίας το οποίο ορίστηκε πρώτη φορά σύμφωνα με τον Nelsen το 1904 . Όπως θα δούμε και ο συντελεστής Spearman's ρ_S είναι μια κανονικοποιημένη αναμενόμενη τιμή και επιπλέον εκφράζει το βαθμό συσχέτισης.

Ο επίσημος ορισμός του συντελεστή Spearman's ρ_S δίνεται από τον παρακάτω ορισμό :

Ορισμός 15 *Έστω οι τ.μ. X, Y με σύζευξη C ο συντελεστής Spearman's ρ_S ή ρ_{SC} ορίζεται ως :*

$$\rho_S = 12 \int\int_{I^2} C(v,z)dvdz - 3 = 12 \int\int_{I^2} vzdC(v,z) - 3 \tag{4.11}$$

Είτε εναλλακτικά $\rho_S = 12 \int\int_{I^2} [C(v,z) - vz]dvdz$

Το μέτρο αυτό επίσης βασίζεται στην πιθανότητα μεταξύ συμφωνίας και ασυμφωνίας . Για την ακρίβεια ξεκινώντας με τρία ανεξάρτητα και όμοια κατανεμημένα τυχαία διανύσματα $(X_1, Y_1), (X_2, Y_2)$ και (X_3, Y_3) με σύζευξη C και ίδια διδιάστατη συνάρτηση κατανομής θα δούμε ότι είναι η διαφορά της τριπλάσιας πιθανότητας συμφωνίας μείον τη πιθανότητα ασυμφωνίας των δύο διανυσμάτων $(X_1, Y_1), (X_2, Y_3)$. Ωστόσο στην περίπτωση του συντελεστή Spearman's ρ_S υπολογισμένη για την ανεξάρτητη περίπτωση είναι :

Θεώρημα 16 *Αν $(X_1, Y_1), (X_2, Y_2)$ και (X_3, Y_3) ανεξάρτητα και όμοια κατανεμημένα τυχαία διανύσματα με με σύζευξη C τότε*

$$\rho_S = 3\mathbb{P}\big[(X_1 - X_2)(Y_1 - Y_3) > 0\big] - \mathbb{P}\big[(X_1 - X_2)(Y_1 - Y_3) < 0\big] \tag{4.12}$$

Ωστόσο αν οι ολοκληρωτικοί μετασχηματισμοί $U_1 = F_1(X), U_2 = F_2(Y)$ ακολουθούν τυπική ομοιόμορφη κατανομή με από κοινού συνάρτηση κατανομής C η σχέση 4.11 του ορισμού 15 ισοδυναμεί με :

$$\rho_S = 12\mathbb{E}[U_1U_2] - 3 = \frac{\mathbb{E}[U_1U_2] - \frac{1}{4}}{\frac{1}{12}}$$

και αφού $\frac{1}{2}$ και $\frac{1}{12}$ είναι η μέση τιμή και η διασπορά της τυπικής ομοιόμορφης κατανομής έπεται ότι :

$$\rho_S = \frac{cov(F_1(X), F_2(Y))}{\sqrt{var(F_1(X))var(F_2(Y))}} \tag{4.13}$$

Αποδεικνύεται και ότι ο συντελεστής Spearman's ρ_S είναι ένα μέτρο συμφωνίας καθώς ικανοποιεί όλες τις ιδιότητες του ορισμού 13. Εστιάζοντας περισσότερο στην δεύτερη ιδιότητα δηλαδή ότι για τον Spearman's ρ_S ισχύει $-1 \leq \rho_S \leq 1$ συνδυάζοντας ωστόσο και το θεώρημα 11 έχουμε ότι το κάτω φράγμα του Spearman's ρ_S επιτυγχάνεται για αντιμονότονες τ.μ. ενώ το άνω φράγμα του για συμμονότονες τ.μ. . Έτσι ,

$$\rho_S = 1 \Leftrightarrow C = C^- \tag{4.14}$$

$$\rho_S = -1 \Leftrightarrow C = C^+ \tag{4.15}$$

Παράδειγμα 12 *Χρησιμοποιώντας τον ορισμό 15 μπορούμε να υπολογίσουμε τον συντελεστή Spearman's ρ_S για τη σύζευξη γινόμενο $C^\perp$:*

$$\begin{aligned}
\rho_{SC^\perp} &= 12\int_0^1\int_0^1 C^\perp(v,z)dvdz - 3 \\
&= 12\int_0^1\int_0^1 vzdvdz - 3 \\
&= 12\Big(\int_0^1 z(\frac{1}{2} - 0)dz\Big) - 3 \\
&= 12\frac{1}{2}(\frac{1}{2} - 0) - 3 \\
&= 3 - 3 = 0
\end{aligned}$$

Όπως στην περίπτωση του συντελεστή Kendall's τ έτσι και στην περίπτωση του συντελεστή Spearman's ρ_S ο συντελεστής για την σύζευξη C με την σύζευξη επιβίωσης $\bar{C}$ συμπίπτουν :

Θεώρημα 17 *Ο συντελεστής Spearman's ρ_S της σύζευξης C με της σύζευξης επιβίωσης $\bar{C}$ συμπίπτουν :*

$$\rho_{SC} = \rho_{S\bar{C}}$$

Αν ωστόσο κάποιος θέλει να υπολογίσει έναν εκτιμητή για τον συντελεστή Spearman's ρ_S όπως ο Gibbons το 1992, έχοντας ένα τυχαίο δείγμα n μεταβλητών

$$(X_i, Y_i) \quad \text{για} \quad i = 1, ..., n$$

αφού ο ρ_S βάσει της σχέσης 4.13 είναι η τάξη συσχέτισης συνδέοντας την τάξη συσχέτισης των δειγματικών τ.μ. :

$$R_i \equiv rank(X_i) \qquad S_i \equiv rank(Y_i)$$

διατάσσοντας σε αύξουσα σειρά προκύπτει ο ακόλουθως δειγματικός συντελεστής Spearman's ρ_S :

$$\hat{\rho_S} = \frac{\sum_{i=1}^{n}(R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^{n}(R_i - \bar{R})^2 \sum_{i=1}^{n}(S_i - \bar{S})^2}} \tag{4.16}$$

Λαμβάνοντας υπόψην ότι οι τάξεις των n δεδομένων είναι οι πρώτοι n ακέραιοι ο παραπάνω τύπος απλοποιείται ως :

$$\hat{\rho_S} = 12\frac{\sum_{i=1}^{n}(R_i - \bar{R})(S_i - \bar{S})}{n(n^2-1)} \tag{4.17}$$

είτε,

$$\hat{\rho_S} = 1 - 6\frac{\sum_{i=1}^{n}(R_i - S_i)^2}{n(n^2-1)} \tag{4.18}$$

Να προσθέσουμε ότι ο παραπάνω δειγματικός εκτιμητής είναι αμερόληπτος.

4.2.3 Η σχέση που συνδέει τον συντελεστή Spearman's ρ_S με τον Kendall's τ

Παρόλο που και οι δύο συντελεστές μετρούν την πιθανότητα συμφωνίας μεταξύ τ.μ. με δοσμένη σύζευξη C οι τιμές των ρ_S και τ είναι συχνά αρκετά διαφορετικοί. Ωστόσο οι Durbin και Stuart το 1951 έδειξαν ότι οι συντελεστές ρ_S και τ για τ.μ. X,Y με σύζευξη C συνέονται με την παρακάτω συναρτησιακή σχέση :

$$\frac{3}{2}\tau - \frac{1}{2} \leq \rho_S \leq \frac{1}{2} + \tau - \frac{1}{2}\tau^2 \qquad \text{για} \quad \tau \geq 0 \tag{4.19}$$

$$-\frac{1}{2} + \tau + \frac{1}{2}\tau^2 \leq \rho_S \leq \frac{3}{2}\tau + \frac{1}{2} \qquad \text{για} \quad \tau < 0 \tag{4.20}$$

(για την απόδειξη των παραπάνω σχέσεων δείτε [DS])

4.2.4 Γραμμικός συντελεστής συσχέτισης $\rho_{X,Y}$ γνωστός και ως συντελεστής Pearson (Pearson product-moment)

Για τ.μ. που ανήκουν στον L^2 η έννοια της συμφωνίας με τη χαλαρή έννοια θα πρέπει να αποτυπώνεται και από την συνδιακύμανση. Επειδή όμως η συνδιακύμανση δεν είναι κανονικοποιημένο ή σχετικό μέτρο παρουσιάζουμε τον συντελεστή Pearson ή γραμμικό συντελεστή συσχέτισης. Δεν αποτελεί μέτρο συμφωνίας καθώς δεν ικανοποιεί ολες τις ιδιότητες του ορισμού 13 παρά μόνο μερικές απο αυτές τις οποίες θα δούμε αναλυτικά παρακάτω.
Πριν δώσουμε τον ορισμό ας ορίσουμε ως $var(X)$ τη διασπορά της τ.μ. X και ας υπενθυμίσουμε ότι ως μή-εκφυλισμένη τ.μ. X ονομάζουμε αυτή την οποία έχει μή μηδενική διασπορά. Ακολουθεί ο ορισμός του γραμμικού συντελεστή συσχέτισης.

Ορισμός 16 *Για μή εκφυλισμένες τ.μ. X, Y που ανήκουν στον L^2 ο γραμμικός συντελεστής συσχέτισης $\rho_{X,Y}$ ορίζεται ως :*

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sqrt{var(X)var(Y)}} \tag{4.21}$$

Θεώρημα 18 *Ο γραμμικός συντελεστής συσχέτισης $\rho_{X,Y}$ ικανοποιεί τις ιδιότητες 1 εως 5 και 7 του ορισμού 13 του μέτρου συμφωνίας $M_{X,Y}$.*

Παρόλα αυτά ο γραμμικός συντελεστής συσχέτισης $\rho_{X,Y}$ δεν ικανοποιεί την ιδιότητα 6 του ορισμού γιαυτο και δεν συγκαταλέγεται σε ένα από τα μέτρα συμφωνίας. Ωστόσο ικανοποιεί τις παρακάτω ιδιότητες οι οποίες τελικά 'παραβιάζουν' τα θεωρήματα 10 και 11 .

Ιδιότητα 1
Ο γραμμικός συντελεστής συσχέτισης $\rho_{X,Y}$ παραμένει αμετάβλητος για γραμμικούς μετασχηματισμούς αύξουσων συναρτήσεων όχι όμως για μή γραμμικούς.
Απόδειξη:
Για την απόδειξη του πρώτου θεωρούμε τους γραμμικούς μετασχηματισμούς $aX + b, cY + d$ για $a, c \in \mathbb{R}^*$ και $b, d \in \mathbb{R}$των τ.μ. X, Y τότε :

$$\begin{aligned}
\rho_{aX+b,cY+d} &= \frac{cov(aX+b, cY+d)}{\sqrt{var(aX+b)var(cY+d)}} \\
&= \frac{ac \cdot cov(X,Y)}{\sqrt{a^2c^2 var(X)var(Y)}} \\
&= \frac{ac \cdot cov(X,Y)}{ac\sqrt{var(X)var(Y)}} \\
&= \frac{cov(X,Y)}{\sqrt{var(X)var(Y)}} \\
&= \rho_{X,Y}
\end{aligned}$$

Για να δείξουμε ότι ο γραμμικός συντελεστής συσχέτισης $\rho_{X,Y}$ δεν είναι αμετάβλητος για μή γραμμικούς αυξουσων μετασχηματισμών ας θεωρήσουμε το επόμενο παράδειγμα .

Παράδειγμα 13 *Ξεκινώντας με δυο τ.μ. X και Y με διδιάστατη συνάρτηση τυπικής κανονικής κατανομής με γραμμικό συντελεστή συσχέτισης $\rho_{X,Y}$ και τους γραμμικούς μετασχηματισμούς τους σύμφωνα με τη συνάρτηση της τυπικής κανονικής κατανομής Φ ,Φ(X) και Φ(Y) , τότε υπολογίζοντας γραμμικό συντελεστή συσχέτισης των Φ(X) και Φ(Y) έχουμε :*

$$\rho_{\Phi(X),\Phi(Y)} = \frac{6}{\pi} arcsin(\frac{\rho_{X,Y}}{2})$$

Ο γραμμικός συντελεστής συσχέτισης $\rho_{X,Y}$ σε αντίθεση με τους προηγούμενους συντελεστές δεν πιάνει τα φράγματα $-1, +1$ απαραίτητα όταν οι τ.μ. X και Y είναι συμμονότονες και αντιμονότονες χωρίς να είναι γραμμικά συσχετισμένες. Επίσημα έχει την επόμενη ιδιότητα.

Ιδιότητα 2

Ο γραμμικός συντελεστής συσχέτισης $\rho_{X,Y}$ είναι φραγμένος

$$\rho_l \leq \rho_{X,Y} \leq \rho_u$$

όπου τα φράγματα ρ_l και ρ_u ορίζονται ως :

$$\rho_l = \frac{\int\int_D (C^-(F_1(x), F_2(y)) - F_1(x)F_2(y))dxdx}{\sqrt{\int_{DomF_1}(x - EX)^2 \int_{DomF_2}(y - EY)^2 dF_2(y)}} \tag{4.22}$$

$$\rho_u = \frac{\int\int_D (C^+(F_1(x), F_2(y)) - F_1(x)F_2(y))dxdx}{\sqrt{\int_{DomF_1}(x - EX)^2 dF_1(x) \int_{DomF_2}(y - EY)^2 dF_2(y)}} \tag{4.23}$$

και επιτυγχάνονται όταν οι τ.μ. X και Y είναι αντιμονότονες και συμμονότονες αντίστοιχα.

Απόδειξη:

Τα φράγματα για τον $\rho_{X,Y}$ από τον τύπο του Hoeffding (1940) για την συνδιακύμανση $cov(X,Y) = \int\int_D (F(x,y) - F_1(x)F_2(y))dxdx$ όπου βάσει του θεωρήματος Sklar γίνεται $cov(X,Y) = \int\int_D (C(F_1(x), F_2(y)) - F_1(x)F_2(y))dxdx$ καθώς και από την ανισότητα *Fréchet* $C^-(v,z) \leq C(v,z) \leq C^+(v,z)$ τότε ισοδύναμα έχουμε :

$$\int\int_D (C^-(F_1(x), F_2(y)) - F_1(x)F_2(y))dxdx \leq cov(X,Y)$$

$$\leq \int\int_D (C^+(F_1(x), F_2(y)) - F_1(x)F_2(y))dxdx$$

και διαιρώντας με την τετραγωνική ρίζα των διασπορών

$$var(X) = \int_{DomF_1} (x - EX)^2 dF_1(x)$$

και της τ.μ. Y αντίστοιχα έπεται το ζητούμενο. □

Ιδιότητα 3

Ο γραμμικός συντελεστής συσχέτισης $\rho_{X,Y}$ για συμμονότονες (ή αντιμονότονες αντίστοιχα)τ.μ. μπορεί να πάρει τιμή διαφορετική του 1 (του -1 αντίστοιχα).

Απόδειξη:

Για να το αποδείξουμε θα θεωρήσουμε ένα παράδειγμα συμμονοτονίας στο οποίο ο γραμμικός συντελεστής συσχέτισης $\rho_{X,Y}$ έχει άνω φράγμα διαφορετικό του 1 . Πριν από αυτό ας παρατηρήσουμε ότι τα φράγματα βάσει της ιδιότητας 2 εξαρτώνταιαπό τις περιθώριες συναρτήσεις. Βάσει αυτής της εξάρτησης μπορούμε εκ των προτέρων να αποτρέψουμε τον συντελεστή να πάρει την τιμή 1 σε απόλυτη τιμή για κάθε ζεύγος συμμονότινων ή αντιμονότονων τ.μ. .

Ας θεωρήσουμε ας παράδειγμα το ακόλουθο βάσει του Wang (1998) :

Παράδειγμα 14 *Έστω ότι οι τ.μ. X, Y ακολουθούν λογαριθμική κανονική κατανομή με παραμέτρους (μ_X,σ_X) , (μ_Y,σ_Y) τα φράγματα παίρνουμε :*

$$\rho_l = \frac{\exp(-\sigma_X \sigma_Y) - 1}{\sqrt{(\exp(\sigma_X^2) - 1)(\exp(\sigma_Y^2) - 1)}} \leq 0 \tag{4.24}$$

$$\rho_l = \frac{\exp(\sigma_X \sigma_Y) - 1}{\sqrt{(\exp(\sigma_X^2) - 1)(\exp(\sigma_Y^2) - 1)}} \geq 0 \tag{4.25}$$

το κάτω φράγμα ρ_l τείνει στο -1 όταν $max(\sigma_X, \sigma_Y) \to 0$ ενώ το άνω φράγμα ρ_u ισούται με 1 αν και μόνο αν $\sigma_X = \sigma_Y$. Ενώ όταν οι τυπικές αποκλίσεις είναι διαφορετικές το διάστημα [ρ_l,ρ_u] είναι διάφορο του [-1,1]. Ακόμα χειρότερα μπορεί να συμβεί το διάστημα [ρ_l,ρ_u] να είναι πολύ ¨μικρό' αν

$$\lim_{max(\sigma_X, \sigma_Y) \to \infty} \rho_l = 0 \tag{4.26}$$

$$\lim_{|\sigma_X - \sigma_Y| \to \infty} \rho_u = 0 \tag{4.27}$$

Ιδιότητα 4

Ο γραμμικός συντελεστής συσχέτισης $\rho_{X,Y} = 0$ δεν συνεπάγεται ότι οι τ.μ. X, Y είναι ανεξάρτητες εκτός και αν είναι Gaussian.

Απόδειξη:

Αν $\rho_{X,Y} = 0$ και οι τ.μ. X, Y είναι Gaussian όπως είναι γνωστό από τη θεωρία πιθανοτήτων οι τ.μ. είναι ανεξάρτητες. Η απόδειξη ότι ο $\rho_{X,Y} = 0$ δεν συνεπάγεται ότι οι τ.μ. X, Y είναι ανεξάρτητες δίνεται μέσω αντιπαραδείγματος το οποίο παραθέτουμε στην απόδειξη της επόμενης ιδιότητας. □

Ιδιότητα 5

Ο γραμμικός συντελεστής συσχέτισης $\rho_{X,Y} = 0$ δεν σημαίνει ότι μια τ.μ. δεν μπορεί να είναι σχεδόν σίγουρα μια συνάρτηση της άλλης.

Απόδειξη:
Η παρακάτω σύζευξη ορίστηκε από τον Nelsen (1999)

$$C(v,z) = \begin{cases} v & ,0 \le v \le \frac{z}{2} \le \frac{1}{2} \\ \frac{z}{2} & ,0 \le \frac{z}{2} \le v \le 1-\frac{z}{2} \\ v+z-1 & ,\frac{1}{2} \le 1-\frac{z}{2} \le v \le 1 \end{cases}$$

Για την δοσμένη σύζευξη η συνδιακύμανση των τ.μ. U_1, U_2 είναι μηδέν επομένως $cov(U_1, U_2) = 0 \Rightarrow \rho_{U_1,U_2} = 0$ (παράδειγμα στο οποίο ο $\rho_{X,Y} = 0$ αλλά οι τ.μ. δεν είναι ανεξάρτητες) αλλά

$$\mathbb{P}(U_2 = 1 - |2U_1 - 1|) = 1$$

δηλαδή οι τ.μ. U_1, U_2 είναι ασυσχέτιστες, αλλά η μία είναι σχεδόν σίγουρα συνάρτηση της άλλης. □

4.2.5 Εξάρτηση από τα άκρα (tail dependence)

Οι περισσότερες έννοιες εξάρτησης είναι σχεδιασμένες για να περιγράψουν πόσο μεγάλες (ή μικρές) τιμές της μιας τ.μ. εμφανίζονται με μεγάλες (ή μικρές) τιμές της άλλης. Μια άλλη τέτοια έννοια είναι η Εξάρτηση από τα άκρα η οποία μετρά την εξάρτηση μεταξύ των τ.μ. στο πάνω δεξιά και κάτω αριστερά τεταρτημόριο του I^2.
Πριν δώσουμε τον παρακάτω ορισμό ας θυμηθούμε τη σύζευξη επιβίωσης $\bar{C}$ και την από κοινού συνάρτηση επιβίωσης για (τυπικά) ομοιόμορφες τ.μ. $\not{C}$ που είδαμε στο τέλος του τρίτου κεφαλαίου. Η σχέση που τις συνέδεε ήταν :

$$\not{C}(v,z) = \bar{C}(1-v, 1-z) = 1 - v - z + C(v,z)$$

$$\text{και} \quad \not{C} = \mathbb{P}(U_1 > v, U_2 > z)$$

Ορισμός 17 *Αν το όριο*

$$\lim_{v \to 1^-} \frac{\not{C}(v,v)}{1-v} = \lambda_U \tag{4.28}$$

υπάρχει και είναι πεπερασμένο. Τότε η σύζευξη C θα λέμε ότι έχει άνω άκρο εξάρτησης αν και μόνο αν $\lambda_u \in (0,1]$ και ότι δεν έχει άνω άκρο εξάρτησης αν και μόνο αν $\lambda_u = 0$. Αντίστοιχα αν το όριο

$$\lim_{v \to 0^+} \frac{C(v,v)}{v} = \lambda_L \tag{4.29}$$

υπάρχει και είναι πεπερασμένο. Τότε η σύζευξη C θα λέμε ότι έχει κάτω άκρο εξάρτησης αν και μόνο αν $\lambda_U \in (0,1]$ και ότι δεν έχει κάτω άκρο εξάρτησης αν και μόνο αν $\lambda_L = 0$

Για να καταλάβουμε καλύτερα τον παραπάνω ορισμό ας δούμε τι συμβολίζει το παραπάνω πηλίκο :

$$\lim_{v \to 1^-} \frac{\bar{C}(v,v)}{1-v} = \frac{\mathbb{P}(U_1 > v, U_2 > v)}{\mathbb{P}(U_2 > v)} = \mathbb{P}(U_1 > v | U_2 > v) = \mathbb{P}(U_2 > v | U_1 > v) \tag{4.30}$$

Ως εκ τούτου το άνω άκρο εξάρτησης είναι τό όριο της δεσμευμένης πιθανότητας :

$$\lambda_U = \lim_{v \to 1^-} = \mathbb{P}(U_1 > v | U_2 > v) = \mathbb{P}(U_2 > v | U_1 > v)$$

Όμοια για κάτω άκρο εξάρτησης έχουμε ,

$$\lambda_L = \lim_{v \to 0^+} = \mathbb{P}(U_1 < v | U_2 < v) = \mathbb{P}(U_2 < v | U_1 < v)$$

Η τιμή του λ_u παριστάνει το όριο της δεσμευμένης πιθανότητας της συνάρτησης κατανομής της Χ να υπερβαίνει το κατώφλι ν δεδομένου ότι η συνάρτηση κατανομής της Υ το κάνει και όταν το ν τείνει στο 1 η τ.μ. υποθέτει άνω ακραίες τιμές. Ανάλογα συμπεράσματα βγαίνουν και για το κάτω άκρο εξάρτησης.
Ας δούμε ένα παράδειγμα υπολογισμού των άκρων εξάρτησης μιας σύζευξης C.

Παράδειγμα 15 *Έστω η σύζευξη* $C(v,z) = pC^+(v,z) + (1-p)C^{\perp}(v,z)$ *τότε* ,

$$\begin{aligned}
\frac{\bar{C}(v,v)}{1-v} &= \frac{1-v-v+C(v,v)}{1-v} \\
&= \frac{1-2v+pmin(v,v)+(1-p)v^2}{1-v} \\
&= \frac{1-2v+pv+(1-p)v^2}{1-v} \\
&= \frac{1-(2-p)v+(1-p)v^2}{1-v}
\end{aligned}$$

Έτσι για το άνω άκρο εξάρτησης έχουμε :

$$\begin{aligned}
\lambda_U &= \lim_{v \to 1^-} \frac{\bar{C}(v,v)}{1-v} \\
&= \frac{1-(2-p)v+(1-p)v^2}{1-v} \\
&=^{DLH} \lim_{v \to 1^-} \frac{-(2-p)+2(1-p)v}{-1} \\
&= \lim_{v \to 1^-} \frac{p-2pv}{-1} = p
\end{aligned}$$

Για το κάτω άκρο εξάρτησης τώρα ,

$$\begin{aligned}\frac{C(v,v)}{v} &= \frac{p\ min(v,v) + (1-p)v^2}{v}\\ &= \frac{pv + (1-p)v^2}{v}\\ &= p + (1-p)v\end{aligned}$$

και

$$\lambda_L = \lim_{v \to 0^+} \frac{C(v,v)}{v} = p$$

Τα άκρα εξάρτησης και οι συντελεστές της με τη σειρά τους προέρχονται από το γεγονός ότι η μέγιστη σύζευξη C^+ παρουσιάζει άκρο εξάρτησης με άνω και κάτω άκρο με τιμή 1. Ενώ η γινόμενο σύζευξη $C^\perp$ καθώς και η ελάχιστη C^- δεν έχουν ούτε άνω ούτε κάτω άκρο εξάρτησης.

Μπορούμε επίσης να παρουσιάσουμε συντελεστές $\bar{\lambda}_U$, $\bar{\lambda}_L$ των άκρων εξάρτησης για τη σύζευξη επιβίωσης $\bar{C}$.

$$\bar{\lambda}_U = \lim_{v \to 1^-} \frac{1 - 2\ v + \bar{C}(v,v)}{1-v}$$

$$\bar{\lambda}_L = \lim_{v \to 0^+} \frac{\bar{C}(v,v)}{v}$$

Αν τα παραπάνω όρια είναι πεπερασμένα τότε ισχύει η παρακάτω ιδιότητα :

Θεώρημα 19 *Έστω $\bar{C}$ η σύζευξη επιβίωσης της σύζεξης C τότε*

$$\bar{\lambda}_U = \lambda_L$$

$$\bar{\lambda}_L = \lambda_U$$

Απόδειξη:
Για το $\bar{\lambda}_U$ έχουμε

$$\begin{aligned}\bar{\lambda}_U &= \lim_{v \to 1^-} \frac{1 - 2\ v + \bar{C}(v,v)}{1-v}\\ &= \frac{C(1-v, 1-v)}{1-v}\\ &= \lim_{v \to 0^+} \frac{C(v,v)}{v}\\ &= \lambda_L\end{aligned}$$

Αντίστοιχα για το $\bar{\lambda}_L$ *έχουμε*

$$\begin{aligned}\bar{\lambda}_L &= \lim_{v\to 0^+} \frac{\bar{C}(v,v)}{v} \\ &= \lim_{v\to 0^+} \frac{2\ v - 1 + C(1-v, 1-v)}{v} \\ &= \lim_{w\to 1^-} \frac{1 - 2\ w + \bar{C}(w,w)}{1-w} \\ &= \lambda_U \quad \square\end{aligned}$$

4.2.6 Θετικά τεταρτημόρια εξάρτησης Θ.Τ.Ε. (positive quadrant dependency PQD)

Η έννοια των θετικών τεταρτημορίων εξάρτησης Θ.Τ.Ε. σύμφωνα με τον Lehmann (1996) μπορεί να εκφραστεί σε όρους σύζευξης ως εξής :

Ορισμός 18 *Οι τ.μ.* X, Y *με σύζευξη* C *θα λέγονται θετικά εξαρτημένες αν και μόνο αν*

$$C(v,z) \geq vz \qquad \text{για κάθε } (v,z) \in I^2 \ .$$

Ένας εναλλακτικός ορισμός χρησιμοποιώντας την συμφωνία διάταξης των συζεύξεων είναι ο εξής :
Οι τ.μ. X,Y θα λέγονται θετικά εξαρτημένες αν και μόνο αν η συνάρτηση σύζευξης τους C είναι μεγαλύτερη επό αυτή της γινόμενο σύζευξης $C^{\perp}$

$$C \succ C^{\perp}$$

Παράδειγμα 16 *Οι τ.μ. με σύζευξη* $C(v,z) = pC^{+}(v,z) + (1-p)C^{\perp}(v,z)$ $p \in I^2$ *είναι θετικά εξαρτημένες αφού σύμφωνα με την ανισότητα* **Fréchet** *1.8 έχουμε* $C^{+} \succ C^{\perp}$ *και συνεπάγεται ότι :*

$$pC^{+} + (1-p)C^{\perp} \succ C^{\perp} \quad \Rightarrow \quad C \succ C^{\perp}$$

Ενώ οι τ.μ. με σύζευξη $C(v,z) = pC^{-}(v,z) + (1-p)C^{\perp}(v,z)$ $p \in I^2$ *δεν είναι θετικά εξαρτημένες αφού* $C^{-} \prec C^{\perp}$ *έτσι συνεπάγεται ότι :*

$$pC^{-} + (1-p)C^{\perp} \succ C^{\perp} \quad \Rightarrow \quad C \succ C^{\perp}$$

Ο εναλλακτικός ορισμός χρησιμοποιώντας συναρτήσεις κατανομής μπορεί να επαναδιατυπωθεί ως :

$$F(x,y) \geq F_1(x)F_2(y) \quad \text{για κάθε } (x,y) \in \mathbb{R}^{*2}$$

θα πρέπει δηλαδή η από κοινού πιθανότητα σε κάθε σημείο να είναι μεγαλύτερη από αυτή της ανεξάρτητης. Τα θετικά τεταρτημόρια εξάρτησης Θ.Τ.Ε. μας διασφαλίζουν ότι οι συντελεστές Spearman's ρ_S , Kendall's τ και ο γραμμικός συντελεστής συσχέτισης $\rho_{X,Y}$ θα είναι μή αρνητικοί αφού για ανεξάρτητες τ.μ. για τις οποίες

γνωρίζουμε ότι $C = C^{\perp}$ μηδενίζονται οι παραπάνω συντελεστές. Επομένως δεδομένου ότι οι παραπάνω συντελεστές ακολουθούν τη συμφωνία διάταξης έχουμε ότι για θετικά εξαρτημένες τ.μ. δεν μπορούν παρά να είναι μή αρνητικοί.
Εφαρμόζωντας τον τύπο του Bayes έχουμε :

$$\begin{aligned}
&F(x,y) \geq F_1(x)F_2(y) \Leftrightarrow \\
&\mathbb{P}\big[X \leq x,\ Y \leq y\big] \geq \mathbb{P}\big[X \leq x,\big]\mathbb{P}\big[Y \leq y\big] \\
&\mathbb{P}\big[X \leq x|Y \leq y\big]\mathbb{P}\big[Y \leq y\big] \geq \mathbb{P}\big[X \leq x,\big]\mathbb{P}\big[Y \leq y\big] \\
&\mathbb{P}\big[X \leq x|Y \leq y\big] \geq \mathbb{P}\big[X \leq x,\big]
\end{aligned}$$

Έτσι η συνθήκη του Lehmann για να είναι οι τ.μ. X,Y θετικά εξαρτημένες μπορεί ενισχυθεί ζητώντας η δεσμευμένη πιθανότητα να είναι μια μή φθίνουσα συνάρτηση του y. Αυτό σημαίνει ότι η πιθανότητα η απόδοση X_t να παίρνει μικρές τιμές δεν αυξάνεται καθώς οι τιμές της απόδοσης Y_t αυξάνονται. Αυτό αντιστοιχεί σε μια συγκεκριμένη μονοτονία στα άκρα. Έτσι θα λέμε ότι η τ.μ. X είναι στο αριστερό άκρο φθίνουσα στην Y (left tail deacreasing LTD) και συμβολίζουμε ως $LTD(X|Y)$ αν

$$\mathbb{P}\big[X \leq x|Y \leq y\big] \ \text{είναι μια μή φθίνουσα συνάρτηση του} \quad y \ \ \text{για κάθε} \quad x$$

Αυτό με τη σειρά του είναι ισοδύναμο με :
για κάθε $v \in [0,1]$η$\frac{C(v,z)}{z}$ να είναι μια μή φθίνουσα συνάρτηση του z ή

$$\frac{\partial C(v,z)}{\partial z} \leq \frac{C(v,z)}{z} \quad \text{σχεδόν για κάθε} \quad z$$

4.3 Παραμετρικές οικογένειες συναρτήσεων σύζευξης

Στην παράγραφο αυτή θα παρουσιάσουμε μερικές από τις πιο γνωστές οικογένειες συναρτήσεων σύζευξης. Θα δώσουμε για κάθε οικογένεια τον ορισμό , την πυκνότητα σύζευξης και την δεσμευμένη κατανομή σαν συνάρτηση της σύζευξης. Θα ονομάζουμε **περιεκτική** (comprehensive) μια οικογένεια εφόσον περιέχει την ελάχιστη , τη γινόμενο και την μέγιστη σύζευξη. Κάθε οικογένεια θα χαρακτηρίζεται από μια παράμετρο ή ένα διάνυσμα παραμέτρων και στις περιπτώσεις που είναι δυνατό θα διασαφηνίσουμε τα μέτρα συμφωνίας που ορίσαμε παραπάνω.

4.3.1 Gaussian σύζευξη

Ο ορισμός της Gaussian σύζευξης γνωστής και ως κανονικής είναι ο εξής :

Ορισμός 19 *Η Gaussian σύζευξη δίνεται από :*

$$C^{GA}(v,z) = \Phi_{\rho_{X,Y}}\big(\Phi^{-1}(v)\ ,\ \Phi^{-1}(z)\big) \tag{4.31}$$

Όπου $\Phi_{\rho_{X,Y}}$ η από κοινού συνάρτηση κατανομής μιας διδιάστατης κανονικής κατανομής με συντελεστή συσχέτισης $\rho_{X,Y}$ και Φ αυτή της μονοδιάστατης αντίστοιχα. Επομένως

$$C^{GA}(v,z) = \\ = \int_{-\infty}^{\Phi^{-1}(v)} \int_{-\infty}^{\Phi^{-1}(z)} \frac{1}{2\pi\sqrt{1-\rho_{X,Y}^2}} e^{\frac{2\rho_{X,Y} st - s^2 - t^2}{2(1-\rho_{X,Y}^2)}} dsdt \tag{4.32}$$

Εφόσον η Gaussian σύζευξη χαρακτηρίζεται από τον γραμμικό συντελεστή συσχέτισης ρ οι συμβολισμοί C^{GA} και C_ρ^{GA} στο εξής θα είναι ισοδύναμοι . Μια ισοδύναμη αναπαράσταση με την 4.32 η οποία έχει αποδειχθεί από το Roncalli το 2002 είναι η :

$$C^{GA}(v,z) = \int_0^v \Phi\left(\frac{\Phi^{-1}(z) - \rho_{X,Y}\Phi^{-1}(t)}{\sqrt{1-\rho_{X,Y}^2}}\right) dt \tag{4.33}$$

Στο κεφάλαιο 3 ορίσαμε την πυκνότητα σύζευξης ως $c(v,z) = \frac{\partial^2 C(v,z)}{\partial v \partial z}$ έτσι η πυκνότητα της παραπάνω συζευξης δίνεται από :

$$c^{GA}(v,z) = \frac{1}{\sqrt{1-\rho_{X,Y}^2}} e^{\frac{j_1^2+j_2^2}{2} + \frac{2\rho_{X,Y} j_1 j_2 - j_1^2 - j_2^2}{2(1-\rho_{X,Y}^2)}} \tag{4.34}$$

όπου $j_1 \equiv \Phi^{-1}(v)$, $j_2 \equiv \Phi^{-1}(z)$
Εφόσον η σύζευξη είναι απολύτως συνεχής ολοκληρώνοντας την πυκνότητα έχουμε τον παρακάτω ισοδύναμο τύπο :

$$C^{GA}(v,z) = \int_0^v \int_0^z \frac{1}{\sqrt{1-\rho_{X,Y}^2}} e^{\frac{x^2+y^2}{2} + \frac{2\rho_{X,Y} xy - x^2 - y^2}{2(1-\rho_{X,Y}^2)}} dsdt \tag{4.35}$$

όπου $x \equiv \Phi^{-1}(s)$, $y \equiv \Phi^{-1}(t)$ *Όσο για τη δεσμευμένη κατανομή μέσω της σύζευξης δίνεται :*

$$C_{2|1}^{GA}(v,z) = \Phi\left(\frac{\Phi^{-1}(z) - \rho_{X,Y}\Phi^{-1}(v)}{\sqrt{1-\rho_{X,Y}^2}}\right)$$

$$\text{όπου} \quad C_{2|1}^{GA}(v,z) = \mathbb{P}\big[U_2 \le z,\ U_1 = v\big]$$

Μια χρήσιμη ιδιότητα της Gaussian σύζευξης όπως θα δούμε στην επόμενη πρόταση είναι ότι δημιουργεί την διμετάβλητη Gaussian συνάρτηση κατανομής , ειδικότερα :

Πρόταση 1 *Η Gaussian σύζευξη δημιουργεί την από κοινού τυπική κανονική συνάρτηση κατανομής μέσω του θεωρήματος Sklar αν και μόνο αν οι περιθώριες*

είναι τυπικές κανονικές.
Απόδειξη:
Η Gaussian σύζευξη υπολογισμένη στο $F_1(x), F_2(y)$ σύμφωνα με τον ορισμό δίνει

$$C^{GA}(F_1(x), F_2(y)) = \int_{-\infty}^{\Phi^{-1}(F_1(x))} \int_{-\infty}^{\Phi^{-1}(F_2(y))} \frac{1}{2\pi\sqrt{1-\rho_{X,Y}^2}} e^{\frac{2\rho_{X,Y}st - s^2 - t^2}{2(1-\rho_{X,Y}^2)}} dsdt$$

επομένως ισχύει η ισοδυναμία

$$C^{GA}(F_1(x), F_2(y)) = \Phi_{\rho_{X,Y}}\big(x\ ,\ y\big) \;\Leftrightarrow\; \Phi^{-1}(F_1(x)) = x \text{ και } \Phi^{-1}(F_2(y)) = y$$

δηλαδή αν και μόνο αν $F_1 = F_2 = \Phi$ οι περιθώριες είναι τυπικές κανονικές . □
Η Gaussian σύζευξη είναι παραμετρικοποιημένη από τον γραμμικό συντελεστή συσχέτισης ο οποίος ακολουθεί τη συμφωνία διάταξης . Σαν συνέπεια αυτού έχουμε ότι η Gaussian σύζευξη είναι ***θετικά διατεταγμένη*** *σε σχέση με την παράμετρο ρ (positively ordered) :*

$$C^{GA}_{\rho=-1} \prec C^{GA}_{\rho<0} \prec C^{GA}_{\rho=0} \prec C^{GA}_{\rho>0} \prec C^{GA}_{\rho=1} \tag{4.36}$$

Επιπλέον εύκολα διαπιστώνει κανείς με τον ορισμό ότι είναι περιεκτική καθώς $C^{GA}_{\rho=-1} = C^-$ και $C^{GA}_{\rho=1} = C^+$ ενώ για ρ=0 $C^{GA}_{\rho=0} = C^\perp$. Όσο για τα μέτρα συμφωνίας χρησιμοποιώντας τους αντίστοιχους ορισμούς οι συντελεστές Kendall's τ και Spearman's ρ_S για την Gaussian σύζευξη δίνονται από :

$$\tau_{C^{GA}} = \frac{2}{\pi}\arcsin\rho \quad \text{και} \quad \rho_{S,C^{GA}} = \frac{6}{\pi}\arcsin\frac{\rho}{2}$$

Ενώ για την εξάρτηση από τα άκρα για τα όρια λ_U, λ_L έχουμε :

$$\lambda_U = \lambda_L = \begin{cases} 0 & \text{ανν } \rho < 1 \\ 1 & \text{ανν } \rho = 1 \end{cases}$$

Δηλαδή έχουν άνω και κάτω άκρο εξάρτησης μόνο για ρ=1 διαφορετικά όχι . Τέλος σύμφωνα με την σχέση 4.36 και τον ορισμό για τις θετικά εξαρτημένες μεταβλητές έχουμε ότι οι τ.μ. X, Y με σύζευξη την Gaussian σύζευξη θα είναι θετικά εξαρτημένες αν ρ≥ 0 .

4.3.2 Student's t σύζευξη

Έστω $t_v : \mathbb{R} \to \mathbb{R}$ η συνάρτηση της Student's t κατανομής με v βαθμούς ελευθερίας β.ε.

$$t_v(x) = \int_{-\infty}^{x} \frac{\Gamma\big(\frac{v+1}{2}\big)}{\sqrt{\pi v}\,\Gamma\big(\frac{v}{2}\big)} \Big(1 + \frac{s^2}{v}\Big)^{-\frac{v+1}{2}} ds$$

όπου Γ η συνάρτηση Euler με $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$. Έστω $\rho \in I$ και $t_{\rho,v}$ η διμετάβλητη κατανομή σε αντιστοιχία με το t_v :

$$t_{\rho,v}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} \frac{1}{2\pi\sqrt{1-\rho^2}} \big(1 + \frac{s^2 + t^2 - 2\rho st}{v(1-\rho^2)}^{-\frac{v+2}{2}}\big) dsdt$$

Ορισμός 20 *Η διμετάβλητη Student's t σύζευξη $T_{\rho,v}$ ορίζεται ως :*

$$\begin{aligned} T_{\rho,v}(v,z) &= t_{\rho,v}\big(t_v^{-1}(v), t_v^{-1}(z)\big) \\ &= \int_{-\infty}^{t_v^{-1}(v)} \int_{-\infty}^{t_v^{-1}(z)} \frac{1}{2\pi\sqrt{1-\rho^2}} \big(1+\frac{s^2+t^2-2\rho st}{v(1-\rho^2)}^{-\frac{v+2}{2}}\big) dsdt \end{aligned} \tag{4.37}$$

Ωστόσο όταν ο αριθμός των βαθμών ελευθερίας αποκλίνει η σύζευξη συγκλίνει στην Gaussian σύζευξη . Ενώ σε σχέση με την Gaussian σύζευξη έχει περισσότερες παρατηρήσεις στα άκρα.
Η πυκνότητα της Student's σύζευξης δίνεται :

$$c^S_{v,\rho}(v,z) = \rho^{-\frac{1}{2}} \frac{\Gamma\big(\frac{v+2}{2}\big)\Gamma\big(\frac{v}{2}\big)}{\Gamma\big(\frac{v+1}{2}\big)^2} \frac{\big(1+\frac{s_1^2+s_2^2-2\rho s_1 s_2}{v(1-\rho^2)}\big)^{-\frac{v+2}{2}}}{\Pi_{j=1}^2\big(1+\frac{s_j^2}{v}\big)^{-\frac{v+2}{2}}}$$

όπου $s_1 = t_v^{-1}(v) s_1 = t_v^{-1}(z)$ και η σύζευξη είναι απολύτως συνεχής.
Όπως είναι γνωστό από τον Roncalli (2002) για δοσμένο ζεύγος τ.μ. (X,Y) με από κοινού Student's t κατανομή η δεσμευμένη κατανομή της

$$\sqrt{\frac{v+1}{v+x^2}} \frac{Y-\rho x}{\sqrt{1-\rho^2}}$$

δεδομένου $X = x$ είναι μια Student's t κατανομή με $v+1$ βαθμούς ελευθερίας , η δεσμευμένη κατανομή μέσω της σύζευξης δίνεται από :

$$C^S_{2|1v,\rho}(v,z) = t_{v+1}\Big(\sqrt{\frac{v+1}{v+t_v^{-1}(v)^2}} \frac{t_v^- 1(z) - \rho t_v^- 1(v)}{\sqrt{1-\rho^2}}\Big)$$

Επομένως μια ισοδύναμη έκφραση για τη διμετάβλητη Student's σύζευξη είναι η εξής

$$T_{\rho,v}(v,z) = \int_0^v t_{v+1}\Big(\sqrt{\frac{v+1}{v+t_v^{-1}(v)^2}} \frac{t_v^- 1(z) - \rho t_v^- 1(v)}{\sqrt{1-\rho^2}}\Big) ds$$

Αν $v > 2$ κάθε μια από τις περιθώριες έχει πεπερασμένη διακύμανση , $\frac{v}{v-2}$ και ο συντελεστής ρ μπορεί να ερμηνευτεί ως γραμμικός συντελεστής συσχέτισης. Η Student's σύζευξη είναι ***θετικά διατεταγμένη*** *σε σχέση με την παράμετρο ρ για δοσμένους βαθμούς ελευθερίας και επιπλέον πιάνει το άνω και κάτω φράγμα αφού*

$$C^S_{v,-1} = C^- \quad \text{και} \quad C^S_{v,1} = C^+$$

Εντούτοις , $C^S_{v,o} \neq C^\perp$ για πεπερασμένο v . Επομένως η Student's σύζευξη δεν είναι περιεκτική . Όσο για τα άκρα εξάρτησης για πεπερεασμένο v έχουμε :

$$\lambda_U = \lambda_L = \begin{cases} > 0 & \text{ανν } \rho > -1 \\ 0 & \text{ανν } \rho = 1 \end{cases}$$

Επομένως δεν έχει άνω και κάτω άκρο εξάρτησης αν ρ=1.

4.3.3 Οικογένεια Fréchet

Ορισμός 21 *Ο Fréchet (1958) εισήγαγε την παρακάτω οικογένεια συζεύξεων 2-παραμέτρων :*

$$\begin{aligned} C^F(v,z) &= p\max(v+z-1,0)+(1-p-q)vz+q\min(v,z) \\ &= pC^- +(1-p-q)C^\perp + qC^+ \end{aligned} \tag{4.38}$$

όπου $p,q\in I \ , \ p+q\leq 1$.

Η πυκνότητα της Fréchet σύζευξης είναι :

$$C^F(v,z)=1-p-q$$

Έπεται ότι η σύζευξη αυτή έχει μια απολύτως συνεχή και μια ιδιόμορφη συνιστώσα εάν τουλάχιστον ένα από τα p και q είναι θετικό . Όσο για τη δεσμευμένη κατανομή μέσω της σύζευξης δίνεται από :

$$C^F_{2|1}(v,z)=\begin{cases} p+(1-p-q)z+q & v+z-1>0\ ,\ v<z \\ p+(1-p-q)z & v+z-1>0\ ,\ v>z \\ (1-p-q)z & v+z-1<0\ ,\ v>z \\ (1-p-q)z+q & v+z-1<0\ ,\ v<z \end{cases}$$

Η οικογένεια των συζεύξεων Fréchet θεωρείται θετικά διατεταγμένη σε σχέση με το q και αρνητικά διατεταγμένη σε σχέση με το p .
Όπως παρατηρούμε από τον ορισμό της Fréchet σύζευξης ,

$$C^F(v,z)=\begin{cases} C^-(v,z) & \text{για}\quad p=1\ ,\ q=0 \\ C^+(v,z) & \text{για}\quad p=0\ ,\ q=1 \\ C^\perp(v,z) & \text{για}\quad p=q=0 \end{cases}$$

Επομένως η οικογένεια των συζεύξεων Fréchet είναι περιεκτική .
Όσο για τα μέτρα συμφωνίας σύμφωνα με τον Nelsen (1991) , οι συντελεστές Kendall's τ^F και Spearman's ρ_S^F σε σχέση με τις παραμέτρους p και q για την οικογένεια των συζεύξεων Fréchet δίνονται από :

$$\tau^F=\frac{(q-p)(2+p+q)}{3}$$

$$\rho_S^F=q-p$$

το οποίο συνεπάγεται ότι

$$\begin{aligned} \tau^F\leq \rho_S^F &\leq -1+\sqrt{1+3\tau^F} \qquad \tau^F\geq 0 \\ 1-\sqrt{1+3\tau^F}\leq \rho_S^F &\leq \tau^F \qquad \tau^F<0 \end{aligned}$$

Τα παραπάνω φράγματα είναι αυστηρότερα από αυτά στις συναρτησιακές σχέσεις 4.19 και 4.20 .
Στην περίπτωση όπου μηδενίσουμε τη μία από τις δύο παραμέτρους η οικογένεια των συζεύξεων Fréchet περιορίζεται στην γνωστή ως ***μικτή σύζευξη*** *(mixture copula) . Σε κάθε περίπτωση θα συμβολίζεται και ισούται με :*

$$\begin{aligned} C_q^M(v,z) &= (1-q)vz + q\min(v,z) = (1-q)C^{\perp} + qC^{+} \\ C_p^M(v,z) &= (1-p)vz + p\max(v+z-1,0) = (1-p)C^{\perp} + pC^{-} \end{aligned} \tag{4.39}$$

Σε αντίθεση με την Fréchet σύζευξη οι παραπάνω συζεύξεις δεν είναι περιεκτικές καθώς :

$$C_q^M(v,z) = \begin{cases} C^{\perp}(v,z) & \text{για} \quad q=0 \\ C^{+}(v,z) & \text{για} \quad q=1 \end{cases}$$

Αντίστοιχα ,

$$C_p^M(v,z) = \begin{cases} C^{-}(v,z) & \text{για} \quad p=0 \\ C^{\perp}(v,z) & \text{για} \quad p=1 \end{cases}$$

Όσο για τους συντελεστές Kendall's τ και Spearman's ρ_S για τη μικτή σύζευξη C_q^M και τη σχέση που τα συνδέει έχουμε :

$$\tau = \frac{q(2+q)}{3} \quad \text{και} \quad \rho_S = q$$

$$\tau = \rho_S \frac{2+\rho_S}{3} \quad \text{και} \quad \rho_S = -1 + \sqrt{1+3\tau}$$

αντίστοιχα για την C_p^M ,

$$\tau = -\frac{p(2+p)}{3} \quad \text{και} \quad \rho_S = -p$$

$$\tau = \rho_S \frac{2-\rho_S}{3} \quad \text{και} \quad \rho_S = 1 - \sqrt{1-3\tau}$$

Τέλος για την εξάρτηση από τα άκρα για την $C_q^M(v,z)$ τα όρια $\lambda_L = \lambda_U = q$ επομένως η μικτή σύζευξη C_q^M παρουσιάζει άνω και κάτω άκρο εξάρτησης όταν $q > 0$. Αντίστοιχα για την μικτή σύζευξη C_p^M τα όρια $\lambda_L = \lambda_U = 0$ ως εκ τούτου δεν παρουσιάζει εξάρτηση από τα άκρα .

4.3.4 Η Κλάση των Archimedean συζεύξεων

Η Κλάση των Archimedean συζεύξεων ονομάστηκε από τον Ling το 1965 αλλά ανακαλύφθηκε από τους Schweizer και Sklar το 1961. Πρίν εφαρμοστούν στη χρηματοοικονομία χρησιμοποιήθηκαν στο πεδίο της αναλογιστικής. Ας δούμε τον ορισμό και μερικές βασικές ιδιότητες .
Ορισμός και βασικές ιδότητες
Για τον ορισμό της Archimedean σύζευξης εισάγουμε την έννοια της γεννήτριας

συνάρτησης ϕ όπου $\phi : [0,1] \longrightarrow (0,+\infty)$ συνεχής , φθίνουσα και κυρτή τέτοια ώστε $\phi(1) = 0$. Θα ονομάζεται αυστηρή γεννήτρια (strict generator) όταν $\phi(o) = +\infty$. Ορίζουμε και την ψευδο-αντίστροφη της ϕ , $\phi^{[-1]}$ ως :

$$\phi^{[-1]}(v) = \begin{cases} \phi^{-1}(v) & \text{για} \quad 0 \leq v \leq \phi(o) \\ 0 & \text{για} \quad \phi(o) \leq v \leq +\infty \end{cases} \tag{4.40}$$

Ωστόσο οι ιδιότητες μιας συνάρτησης με την αντίστροφη της ισχύουν και για την ψευδο-αντίστροφη $\phi^{[-1]}$ με την ϕ . Πιο συγκεκριμένα η σύνθεση της $\phi^{[-1]}$ με την ϕ μας δίνει την ταυτοτική και το πεδίο ορισμού της $\phi^{[-1]}$ θα είναι το σύνολο τιμών της ϕ , έτσι

$$\phi^{[-1]}\big(\phi(v)\big) = v \quad \forall v \in [0,1] \quad \text{και} \quad \phi^{[-1]} : (0,+\infty) \longrightarrow [0,1]$$

Ας σημειώσουμε ότι η ϕ είναι αυστηρή γεννήτρια τότε $\phi^{[-1]} = \phi$. Μπορούμε να δώσουμε τον ορισμό της Archimedean σύζευξης τώρα .

Ορισμός 22 *Έστω γεννήτρια συνάρτηση ϕ και η ψευδο-αντίστροφη της $\phi^{[-1]}$, η Archimedean σύζευξη C^A ορίζεται ως*

$$C^A(v,z) = \phi^{[-1]}\big(\phi(v) + \phi(z)\big) \tag{4.41}$$

αν η γεννήτρια συνάρτηση ϕ είναι αυστηρή τότε η παραπάνω σύζευξη ονομάζεται αυστηρή Archimedean σύζευξη .

Ας θυμηθούμε τον ορισμό του μετασχηματισμού Laplace,

Ορισμός 23 *Ο μετασχηματισμός Laplace μιας θετικής τ.μ. γ με συνάρτηση κατανομής F_γ ορίζεται ως*

$$\tau(s) = E_\gamma(e^{-s\gamma}) = \int_0^{+} \infty e^{-st} dF_\gamma(t)$$

Ο αντίστροφος μετασχηματισμός Laplace δίνει αυστηρά γεννήτριες , έτσι για παράγουμε Archimedean συζεύξεις είναι επαρκές να ξεκινήσουμε από μια κλάση τέτοιων μετασχηματισμών .
Ιδιότητες της Archimedean σύζευξης

- *είναι συμμετρικές δηλαδή $C^A(v,z) = C^A(z,v)$*
- *είναι συσχετισμένες (associative) δηλαδή*

$$C^A\big(C^A(v,z),u\big) = C^A\big(v, C^A(z,u)\big) \quad \forall (v,z,u) \in [0,1]^3$$

Όσο για την πυκνότητα της Archimedean σύζευξης δίνεται από :

$$c^A(v,z) = \frac{-\phi''\big(C(v,z)\big)\phi'(v)\phi'(z)}{\big(\phi'(C(v,z))\big)^3}$$

Archimedean σύζευξη και μέτρα συμφωνίας

Οι Genest και Mackay το 1986 απέδειξαν ότι ο συντελεστής Kendall's τ για την Archimedean σύζευξη δίνεται από

$$\tau^A = 4\int_0^1 \frac{\phi(v)}{\phi'(v)}dv + 1 \tag{4.42}$$

όπου η $\phi'(v)$ υπάρχει σχεδόν παντού αφού η γεννήτρια είναι κυρτή . Επιπλέον οι Genest και Mackay έδειξαν ότι για δοσμένη συνθήκη των γεννητριών ϕ_1, ϕ_2 δύο Archimedean συζεύξεων C_1, C_2 η διάταξη των συζεύξεων παραμένει ίδια με αυτή των αντίστοιχων συντελεστών τους ,Kendall's τ και Spearman's ρ_S δηλαδή

$$C_1^A \prec C_2^A \Leftrightarrow \tau_{C_1^A} \leq \tau_{C_2^A} \tag{4.43}$$

ισοδύναμα

$$C_1^A \prec C_2^A \Leftrightarrow \rho_{SC_1^A} \leq \rho_{SC_2^A} \tag{4.44}$$

Επομένως ένα καλό χαρακτηριστικό δύο Archimedean συζεύξεων είναι ότι μπορούν να διαταχθούν εύκολα συγκρίνοντας τους αντίστοιχους συντελεστές . Όσο για την εξάρτηση από τα άκρα έχουμε το ακόλουθο θεώρημα το οποίο αποδείχθηκε από τον Joe το 1997 .

Θεώρημα 20 *Έστω ϕ μια αυστηρή γεννήτρια τέτοια ώστε η $\phi^{[-1]}$ στην κλάση μετασχηματισμών Laplace σχεδόν σίγουρα αυστηρά θετικών τ.μ. . Αν το $\phi'(0)$ είναι πεπερασμένο και διάφορο του μηδενός τότε η σύζευξη*

$$C^A(v,z) = \phi^{-1}\big(\phi(v) + \phi(z)\big)$$

δεν έχει άνω άκρα εξάρτησης ,ενώ αν έχει άνω άκρα εξάρτησης τότε $\frac{1}{\phi'(0)} = -\infty$ και ο συντελεστής για το άνω άκρο εξάρτησης είναι

$$\lambda_U = 2 - 2\lim_{s\to 0^+}\frac{\phi'(s)}{\phi'(2s)}$$

Αντίστοιχα ο συντελεστής για το κάτω άκρο εξάρτησης είναι

$$\lambda_L = 2\lim_{s\to +\infty}\frac{\phi'(s)}{\phi'(2s)}$$

Όσο για τα Θ.Τ.Ε. η σχέση που συνδέει την εξαρτησία με τις Archimedean συζεύξεις βασίζεται στην παρακάτω ιδέα .

Ορισμός 24 *Μια συνάρτηση $h : \mathbb{R} \to \mathbb{R}$ θα λέγεται πλήρως μονότονη στο διάστημα J αν είναι C^∞ και αν στο εσωτερικό του J οι παράγωγοι εναλλάσουν το πρόσημο τους ,δηλαδή*

$$(-1^n)\frac{d^n h(t)}{dt^n} \geq 0 \quad \text{για} \quad n = 0, 1, 2, 3...$$

Ειδικότερα παρατηρούμε ότι μια πλήρως μονότονη συνάρτηση πρέπει να είναι μή-αρνητική ($h(t) \geq 0$) και μή-αύξουσα ($h'(t) \leq 0$) .

Έχοντας ορίσει την πλήρως μονότονη συνάρτηση έχουμε το εξής θεώρημα .

Θεώρημα 21 *Αν η αντίστροφη μιας αυστηρής γεννήτριας είναι πλήρως μονότονη τότε η αντίστοιχη σύζευξη είναι θετικά εξαρτημένη , επομένως*

$$\phi^{-1}\big(\phi(v) + \phi(z)\big) \geq vz$$

Οικογένειες Archimedean συζεύξεων μιας παραμέτρου
Μεταξύ των Archimedean συζεύξεων θα παρουσιάσουμε κάποιες από τις πιο γνωστές οικογένειες ιο οποίες παράγονται από κάποια γεννήτρια συνάρτηση $\phi_\alpha(t)$ είσαγοντας μια νέα παράμετρο α όπου $\alpha \in \mathbb{R}$. Επιλέγοντας τη γεννήτρια θα προκύπτει και η αντίστοιχη οικογένεια .

Οικογένεια Gumbel
Η οικογένεια Gumbel γνώστη και ως Gumbel-Hougaard παράγεται από τη γεννήτρια συνάρτηση

$$\phi_\alpha(t) = \big(-lnt\big)^\alpha \quad , \quad \alpha \in [1, +\infty]$$

με συνάρτηση σύζευξης

$$C(v,z) = e^{-[(-lnv)^\alpha + (-lnz)^\alpha]^{1/\alpha}}$$

Ωστόσο με κατάλληλη επιλογή α περιέχει μόνο την σύζευξη γινόμενο και την μέγιστη σύζευξη,

$$C(v,z) = \begin{cases} C^\perp & \text{για} \quad \alpha = 1 \\ C^+ & \text{για} \quad \alpha \to +\infty \end{cases}$$

Επομένως δεν είναι περιεκτική εφόσον δεν περιέχει και την ελάχιστη σύζευξη. Επιπλέον η οικογένεια Gumbel παρουσιάζει ανεξαρτησία όταν η παράμετρος α πιάνει την ελάχιστη τιμή της ένω έχει θετική εξάρτηση όταν η παράμετρος τείνει να πιάσει τη μέγιστη τιμή της . Χρησιμοποιώντας την σχέση 4.42 έχουμε ότι ο συντελεστής Kendall's τ δίνεται από την σχέση $\tau = 1 - \alpha^{-1}$ ενώ για τον συντελεστή Spearman's ρ_S δεν υπαρχει κάποιος κλειστός τύπος .
Όσο για τα άκρα εξάρτησης εφαρμόζοντας το θεώρημα 20 διαπιστώνουμε ότι η οικογένεια Gumbel έχει άνω άκρο εξάρτησης με

$$\lambda_U = 2 - 2^{\frac{1}{\alpha}}$$

Τέλος να προσθέσουμε ότι η διάταξη των συζεύξεων της οικογένειας Gumbel μπορεί να επιτευχθεί χρησιμοποιώντας όχι μόνο τις σχέσεις 4.43 , 4.44 αλλα συγκρίνοντας τις παραμέτρους α της γεννήτριας από την οποία παράγονται καθώς ακολουθούν την ίδια διάταξη δηλαδή

$$C_1 \prec C_2 \Leftrightarrow \quad \alpha_1 \leq \alpha_2$$

όπου α_i *η παράμετρος που αντιστοιχεί στην σύζευξη* C_i .
Οικογένεια Clayton
Η οικογένεια Clayton παράγεται από την γεννήτρια συνάρτηση

$$\phi_\alpha(t) = \frac{1}{\alpha}\big(t^{-\alpha} - 1\big) \quad , \quad \alpha \in [-1, 0) \cup (0, +\infty)$$

με συνάρτηση σύζευξης

$$C(v, z) = \max\big[\big(v^{-\alpha} + z^{-\alpha} - 1\big)^{-1/\alpha}, 0\big]$$

Ωστόσο με κατάλληλη επιλογή α *περιέχει την γινόμενο , τη μέγιστη και την ελάχιστη σύζευξη,*

$$C(v, z) = \begin{cases} C^{\perp} & \text{για} \quad \alpha = 0 \\ C^{+} & \text{για} \quad \alpha \to +\infty \\ C^{-} & \text{για} \quad \alpha \to -\infty \end{cases}$$

Επομένως είναι περιεκτική . Επιπλέον η οικογένεια Clayton παρουσιάζει όλα τα είδη εξάρτησης . Χρησιμοποιώντας την σχέση 4.42 έχουμε ότι ο συντελεστής Kendall's τ *δίνεται από την σχέση* $\tau = \frac{\alpha}{\alpha+2}$ *ωστόσο για τον συντελεστή Spearman's* ρ_S *υπάρχει κάποια πολύπλοκη σχέση .*
Όσο για τα άκρα εξάρτησης εφαρμόζοντας το θεώρημα 20 διαπιστώνουμε ότι η οικογένεια Clayton έχει κάτω άκρο εξάρτησης για $\alpha > 0$ *με*

$$\lambda_L = 2^{-\frac{1}{\alpha}}$$

Τέλος να προσθέσουμε ότι η διάταξη των συζεύξεων της οικογένειας Clayton μπορεί να επιτευχθεί συγκρίνοντας τις παραμέτρους α *της γεννήτριας από την οποία παράγονται όπως στην περίπτωση της οικογένειας Gumbel* δηλαδή

$$C_1 \prec C_2 \Leftrightarrow \quad \alpha_1 \leq \alpha_2$$

όπου α_i *η παράμετρος που αντιστοιχεί στην σύζευξη* C_i .
Οικογένεια Frank
Η οικογένεια Frank παράγεται από την γεννήτρια συνάρτηση

$$\phi_\alpha(t) = -ln\frac{e^{-\alpha t} - 1}{e^{-\alpha} - 1} \quad , \quad \alpha \in (-\infty, 0) \cup (0, +\infty)$$

με συνάρτηση σύζευξης

$$C(v, z) = -\frac{1}{\alpha}ln\big(1 + \frac{(e^{-av} - 1)(e^{-az} - 1)}{e^{-a} - 1}\big)$$

Ωστόσο με κατάλληλη επιλογή α *περιέχει την γινόμενο , τη μέγιστη και την ελάχιστη σύζευξη,*

$$C(v, z) = \begin{cases} C^{\perp} & \text{για} \quad \alpha = 0 \\ C^{+} & \text{για} \quad \alpha \to +\infty \\ C^{-} & \text{για} \quad \alpha \to -\infty \end{cases}$$

Επομένως είναι περιεκτική . Επιπλέον η οικογένεια Frank παρουσιάζει όλα τα είδη εξάρτησης . Να προσθέσουμε ότι η διάταξη των συζεύξεων της οικογένειας Frank μπορεί να επιτευχθεί συγκρίνοντας τις παραμέτρους α της γεννήτριας από την οποία παράγονται όπως στην περίπτωση των οικογενειών Gumbel και Clayton δηλαδή

$$C_1 \prec C_2 \Leftrightarrow \quad \alpha_1 \leq \alpha_2$$

όπου α_i η παράμετρος που αντιστοιχεί στην σύζευξη C_i .
Όσο για τα άκρα εξάρτησης εφαρμόζοντας το θεώρημα 20 διαπιστώνουμε ότι η οικογένεια Frank δεν έχει άκρα εξάρτησης .
Πριν δούμε τις σχέσεις των συντελεστών Kendall's τ και Spearman's ρ_S ας ορίσουμε τις συναρτήσεις Debye οι οποίες χρειάζονται για τον υπολογισμό τους .

$$D_k(\alpha) = \frac{k}{\alpha^k} \int_0^{\alpha} \frac{t^k}{e^t - 1} dt \quad , \; k = 1, 2$$

για τις οποίες ισχύει επιπλέον ότι :

$$D_k(-\alpha) = D_k(\alpha) + \frac{k\alpha}{k+1}$$

Επομένως χρησιμοποιώντας την σχέση 4.42 έχουμε ότι ο συντελεστής Kendall's τ δίνεται από

$$\tau = 1 + 4\frac{D_1(\alpha) - 1}{\alpha}$$

Ενώ ο συντελεστής Spearman's ρ_S δίνεται από :

$$\rho_s = 1 - 12\frac{D_2(-\alpha) - D_2(-\alpha)}{\alpha}$$

4.3.5 Marshall-Olkin σύζευξη

Ορισμός 25 *Η Marshall-Olkin οικογένεια χαρακτηρίζεται από δύο παραμέτρους m,n και ορίζεται ως :*

$$C^{MO}(v,z) = \min(v^{1-m}z, vz^{1-n}) = \begin{cases} v^{1-m}z & \text{για} \quad v^m \geq z^n \\ vz^{1-n} & \text{για} \quad v^m < z^n \end{cases} \tag{4.45}$$

όπου $m, n \in [0, 1]$

Η πυκνότητα της Marshall-Olkin σύζευξης δίνεται από

$$c^{MO}(v,z) = \begin{cases} (1-m)v^{-m} & \text{για} \quad v^m > z^n \\ (1-n)z^{-n} & \text{για} \quad v^m < z^n \end{cases} \tag{4.46}$$

Η δεσμευμένη κατανομή μέσω της σύζευξης δίνεται από

$$C^{MO}_{2|1}(v,z) = \begin{cases} (1-m)v^{-m}z & \text{για} \quad v^m > z^n \\ z^{1-n} & \text{για} \quad v^m < z^n \end{cases} \tag{4.47}$$

Ωστόσο η Marshall-Olkin σύζευξη δίνοντας τις παρακάτω τιμές στις παραμέτρους m,n δίνει

$$C^{MO}(v,z) = \begin{cases} C^{\perp}(v,z) & \text{για} \quad m=0 \ \text{ ή } \ n=0 \\ C^{-}(v,z) & \text{για} \quad m=n=1 \end{cases} \tag{4.48}$$

Επομένως δεν είναι περιεκτική εφόσον δεν περιλαμβάνει και τη μέγιστη σύζευξη C^{+}.

Οι συντελεστές Kendall's τ *Spearman's* ρ_S *για την Marshall-Olkin σύζευξη σε σχέση με τις παραμέτρους m,n δίνονται από τις σχέσεις*

$$\tau^{MO} = \frac{mn}{m - mn + n} \quad \text{και} \quad \rho^{MO} = \frac{3mn}{2m + 2n - mn}$$

Βιβλιογραφία

[CLV] Cherubini U.,Luciano E.,Vecchiato W.: Copula Methods in Finance John Wiley Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England,(2004) (σελ. 49-93)

[Bl] Blomqvist, N.: "The concept of comonotonicity in actuarial science and finance: theory" Ann. Math. Statist., 21,(1950) (σελ. 593-600)

[Ne] Nelsen, R.B.: "Copulas and association" in G. Dall"Aglio, S. Kotz G. Salinetti (eds), Advances in Probability Distributions with Given Marginals. Kluwer Academic Publishers, Dordrecht, pp.,(1991) (σελ. 51-74)

[Ne] Roger B. Nelsen: An Introduction to Copulas Lewis Clark College, MSC 110 ,Portland, OR 97219-7899 USA,(2006)

[DG] Duffie, D. Garleanu, N.: "Risk and valuation of collateralized debt obligations" Finan. Anal.J., 57 (1),(2001) (σελ. 41-59)

[DDGKV] Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R. Vyncke, D.: On a measure of dependence between two random variables" Insur.: Math. Econ. 31,(2002) (σελ. 3-33)

[DS] Durbin, J. Stuart, A.: "Inversions and rank correlations", J. Roy. Statist. Soc., Series B, 2,(1951) (σελ. 303-309)

[Sc] Scarsini, M.: "On measures of concordance", Stochastica, 8, ,(1984) (σελ. 201-218)