

ΕΙΔΙΚΑ ΘΕΜΑΤΑ ΔΙΑΧΕΙΡΙΣΗΣ ΚΙΝΔΥΝΟΥ

Συσχέτιση (Correlation) - Copulas

Σημασία της μέτρησης της συσχέτισης

- Έστω μία εταιρεία που είναι εκτεθειμένη σε δύο μεταβλητές της αγοράς.
- Πιθανή αύξηση των 2 μεταβλητών θα οδηγήσει σε ζημιά για την εταιρία.
- Εάν οι αλλαγές των τιμών των 2 μεταβλητών έχουν υψηλή συσχέτιση, η εταιρία θα έχει πολύ μεγάλη έκθεση στις μεταβολές τους.
- Εάν οι αλλαγές των τιμών των 2 μεταβλητών έχουν αρνητική συσχέτιση, η εταιρία θα έχει μικρή έκθεση στις μεταβολές τους.
- Για την σωστή και ολοκληρωμένη μέτρηση του κινδύνου ένας διαχειριστής θα πρέπει να παρακολουθεί και τις συσχετίσεις μεταξύ των μεταβλητών στις οποίες υπάρχει έκθεση.

Συσχέτιση (Correlation): Ορισμός

- Η συσχέτιση δύο μεταβλητών V_1 και V_2 δίνεται από τον παρακάτω τύπο:

$$\rho = \frac{E(V_1V_2) - E(V_1)E(V_2)}{SD(V_1)SD(V_2)}$$

- Όπου $E(\cdot)$ η αναμενόμενη τιμή και $SD(\cdot)$ η τυπική απόκλιση.
- Εάν δεν υπάρχει συσχέτιση μεταξύ των 2 μεταβλητών:

$$E(V_1V_2) = E(V_1)E(V_2) \text{ και } \rho = 0$$

- Εάν $V_1 = V_2$ τότε ο αριθμητής και ο παρονομαστής του τύπου της συσχέτισης θα είναι ίσοι, άρα $\rho = 1$, κάτι αναμενόμενο.

Συνδιακύμανση (Covariance): Ορισμός

- Ο τύπος της συνδιακύμανσης μεταξύ δύο μεταβλητών V_1 και V_2 :

$$\text{cov}(V_1, V_2) = E(V_1 V_2) - E(V_1)E(V_2)$$

- Όπου $E(\cdot)$ η αναμενόμενη τιμή.
- Ο τύπος της συσχέτισης μπορεί να γραφεί και ως:

$$\rho = \frac{\text{cov}(V_1, V_2)}{SD(V_1)SD(V_2)}$$

- Παρόλο που διαισθητικά είναι πιο εύκολα κατανοητή η έννοια της συσχέτισης (ρ), η συνδιακύμανση χρησιμοποιείται σε μεγάλο βαθμό στα μοντέλα.

Συνδιακύμανση (Covariance): Ορισμός

- Η (δειγματική) συνδιακύμανση μεταξύ δύο μεταβλητών V_1 και V_2 δίνεται από τον παρακάτω τύπο (χρησιμοποιώντας m παρατηρήσεις):

$$\text{cov}(V_1, V_2) = \frac{\sum_{i=1}^m (V_{1,i} - \bar{V}_1)(V_{2,i} - \bar{V}_2)}{m-1}$$

Συνδιακύμανση (Covariance): Ορισμός

- Δύο μεταβλητές ορίζονται ως στατιστικά ανεξάρτητες, όταν η γνώση για την μία από αυτές δεν επηρεάζει την κατανομή πιθανοτήτων της άλλης.
- Αυτό εκφράζεται μαθηματικά ως εξής:

$$f(V_2|V_1 = x) = f(V_2)$$

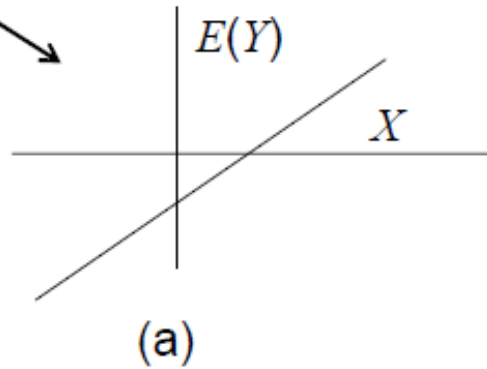
- Όπου $f(\cdot)$ η συνάρτηση πυκνότητας πιθανότητας.
- Εάν η συσχέτιση μεταξύ των 2 μεταβλητών είναι 0, αυτό σημαίνει ότι δεν υπάρχει εξάρτηση μεταξύ τους;
- Όχι απαραίτητα!

Συσχέτιση Vs Εξάρτησης (παράδειγμα)

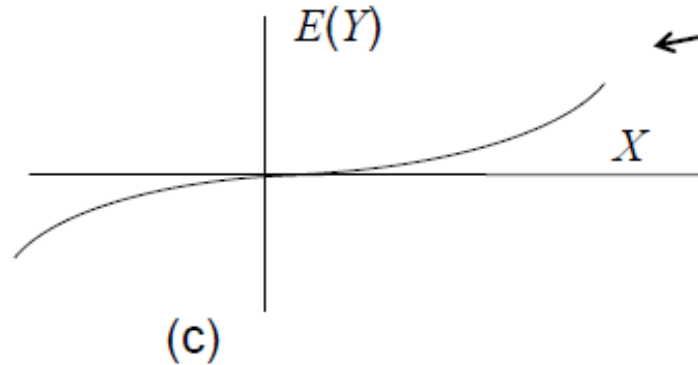
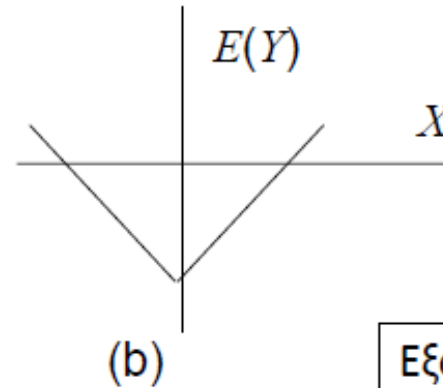
- Αν υποθέσουμε ότι υπάρχουν 3 ισοπίθανες τιμές για την μεταβλητή V_1 : $-1, 0, \text{ ή } +1$.
- Και εάν επιπλέον υποθέσουμε ότι όταν $V_1 = -1$ ή $+1$, τότε $V_2 = 1$ και όταν $V_1 = 0$, τότε και $V_2 = 0$.
- Σε αυτή την περίπτωση:
 - Υπάρχει ξεκάθαρα εξάρτηση μεταξύ των δύο μεταβλητών. Όταν γνωρίζουμε την τιμή της V_1 γνωρίζουμε με βεβαιότητα την τιμή της V_2 , ενώ όταν γνωρίζουμε την V_2 , μας δίνει πληροφορία για την κατανομή πιθανοτήτων της V_1 .
 - Ο συντελεστής συσχέτισης των δύο μεταβλητών είναι ίσος με το μηδέν!
- Ο συντελεστής συσχέτισης μετράει έναν συγκεκριμένο τύπο εξάρτησης, την γραμμική εξάρτηση.
- Υπάρχουν πολλοί διαφορετικοί τύποι εξαρτήσεων, πολλοί από τους οποίους έχουν εφαρμογή και στην μέτρηση του κινδύνου.

Συσχέτιση Vs Εξάρτησης (παράδειγμα)

Γραμμική εξάρτηση



V-shaped εξάρτηση, όπως του παραδείγματος



Εξάρτηση που συναντάμε συχνά σε χρημ/κές μεταβλητές. Όταν η X παίρνει ακραίες τιμές, προκύπτει εξάρτηση με την Y .

Μέτρηση της Συσχέτισης

- Όπως και στην περίπτωση της μεταβλητότητας και της διακύμανσης, έτσι και για την συσχέτιση και την συνδιακύμανση υπάρχουν εναλλακτικοί τρόποι μέτρησης.
- Έστω ότι οι αποδόσεις των μεταβλητών X και Y δίνονται από τους ακόλουθους τύπους: $x_i = (X_i - X_{i-1}) / X_{i-1}$ και $y_i = (Y_i - Y_{i-1}) / Y_{i-1}$.
- Η συνδιακύμανση την ημέρα n : $\text{COV}_n = E(x_n y_n) - E(x_n)E(y_n)$.
- Συχνά γίνεται η απλουστευτική υπόθεση ότι η αναμενόμενη ημερήσια απόδοση είναι ίση με 0, συνεπώς η συνδιακύμανση μετράται ως:
$$\text{COV}_n = E(x_n y_n).$$
- Χρησιμοποιώντας ίση στάθμιση σε έναν αριθμό m τελευταίων παρατηρήσεων, η συνδιακύμανση υπολογίζεται ως:

$$\text{COV}_n = \frac{1}{m} \sum_{i=1}^m x_{n-1} y_{n-1}$$

Μέτρηση της Συσχέτισης (συνέχεια)

- Με παρόμοια στάθμιση μπορούμε να υπολογίσουμε και τις διακυμάνσεις των 2 μεταβλητών:

$$\text{var}_{x,n} = \frac{1}{m} \sum_{i=1}^m x_{n-1}^2 \quad \text{var}_{y,n} = \frac{1}{m} \sum_{i=1}^m y_{n-1}^2$$

- Έχοντας εκτιμήσει την συνδιακύμανση και τις επιμέρους διακυμάνσεις των δύο μεταβλητών, μπορούμε να εκτιμήσουμε και την συσχέτιση:

$$\rho = \frac{\text{cov}_n}{\sqrt{\text{var}_{x,n} \text{var}_{y,n}}}$$

Μέτρηση της Συσχέτισης με το EWMA model

- Στην πράξη, όπως και στην περίπτωση της μεταβλητότητας, οι πρόσφατες παρελθοντικές παρατηρήσεις παρέχουν σημαντικότερη πληροφορία σε σχέση με τις πολύ παλαιότερες παρατηρήσεις.
- Οι διαχειριστές κινδύνου μπορούν να υπολογίσουν την συνδιακύμανση (άρα και την συσχέτιση) δίνοντας διαφορετικές σταθμίσεις αναλόγως την χρονική απόσταση από το παρόν.

- Το μοντέλο EWMA για συνδιακύμανση:

$$\text{COV}_n = \lambda \text{COV}_{n-1} + (1 - \lambda)x_{n-1}y_{n-1}$$

- Όσο πιο υψηλό είναι το «λ» τόσο μεγαλύτερη βαρύτητα δίνουμε στις πιο παλιές παρατηρήσεις και τόσο μικρότερη βαρύτητα στις πρόσφατες (και αντίστροφα).

Μέτρηση της Συσχέτισης με το EWMA model (παράδειγμα)

- Έστω $\lambda=0.95$ και η συσχέτιση μεταξύ X και Y , την ημέρα $n-1$, $\rho_{xy}=0.6$.
- Έστω επίσης ότι οι μεταβλητότητες της X και της Y την ημέρα $n-1$ είναι 1% και 2%, αντίστοιχα.
- Από την σχέση συσχέτισης και συνδιακύμανσης, η συνδιακύμανση την ημέρα $n-1$, είναι ίση με: $\text{COV}_{n-1}=0.6 \times 0.01 \times 0.02=0.00012$.
- Έστω επίσης ότι οι ποσοστιαίες αλλαγές στις τιμές των X και Y , την ημέρα $n-1$, είναι 0.5% και 2.5%, αντίστοιχα.
- Οι διακυμάνσεις και η συνδιακύμανση των X και Y αναθεωρούνται βάση του EWMA ως εξής:

$$\sigma_{x,n}^2 = 0.95 \times 0.01^2 + 0.05 \times 0.005^2 = 0.00009625$$

$$\sigma_{y,n}^2 = 0.95 \times 0.02^2 + 0.05 \times 0.025^2 = 0.00041125$$

$$\text{COV}_n = 0.95 \times 0.00012 + 0.05 \times 0.005 \times 0.025 = 0.00012025$$

Μέτρηση της Συσχέτισης με το EWMA model (παράδειγμα - συνέχεια)

- Η νέα μεταβλητότητα της X είναι 0.981%, ενώ η νέα μεταβλητότητα της Y είναι 2.028% (οι τετραγωνικές ρίζες των νέων διακυμάνσεων).
- Ο νέος συντελεστής συσχέτισης είναι:

$$\rho = \frac{0.00012025}{0.00981 \times 0.02028}$$

- Η μεταβλητότητα της X μειώθηκε, της Y αυξήθηκε, ενώ αυξήθηκε και η συσχέτισή τους.

Μέτρηση της Συσχέτισης με το GARCH model

- Ένα εναλλακτικό μοντέλο με το οποίο μπορούμε να «ενημερώνουμε» τις εκτιμήσεις μας για την συνδιακύμανση (άρα και την συσχέτιση) είναι το GARCH.
- Η διαφορά από το EWMA είναι όπως και στην περίπτωση της διακύμανσης, ότι δίνει στάθμιση και στην μακροχρόνια μέση συνδιακύμανση μεταξύ των μεταβλητών.
- Το μοντέλο GARCH για συνδιακύμανση:

$$\text{COV}_n = \omega + \alpha x_{n-1} y_{n-1} + \beta \text{COV}_{n-1}$$

- Η χρήση του γίνεται με αντίστοιχο τρόπο του GARCH για την διακύμανση και του EWMA για την συνδιακύμανση.
- Μπορεί να γίνει χρήση του GARCH για πρόβλεψη μελλοντικών συνδιακυμάνσεων.

Θέματα Εκτίμησης του πίνακα Διακυμάνσεων - Συνδιακυμάνσεων

- Ο πίνακας αυτός περιέχει στην διαγώνιο τις διακυμάνσεις των μεταβλητών και στα μη-διαγώνια στοιχεία τις αντίστοιχες συνδιακυμάνσεις μεταξύ των μεταβλητών.
- Για να είναι εσωτερικά συνεπής ένας πίνακας $(N \times N)$, θα πρέπει να τηρείται η συνθήκη: $w^T \Omega w \geq 0$ για οποιοδήποτε διάνυσμα w $(N \times 1)$ (ο πίνακας τότε χαρακτηρίζεται ως positive – semidefinite).
- Για να εξασφαλίσουμε ότι τηρείται η παραπάνω συνθήκη, θα πρέπει να προσέχουμε οι διακυμάνσεις και οι συνδιακυμάνσεις να υπολογίζονται με συνεπή τρόπο, δηλαδή με τον ίδιο αριθμό παρατηρήσεων (m).

Θέματα Εκτίμησης του πίνακα Διακυμάνσεων – Συνδιακυμάνσεων (παράδειγμα)

- Ο παρακάτω πίνακας δεν είναι εσωτερικά συνεπής:

$$\begin{pmatrix} 1 & 0 & 0.9 \\ 0 & 1 & 0.9 \\ 0.9 & 0.9 & 1 \end{pmatrix}$$

- Η διακύμανση της κάθε μίας μεταβλητής είναι ίση με 1, άρα η συνδιακυμάνσεις είναι ίσες με τις συσχετίσεις.
- Βλέπουμε ότι:
 - Η πρώτη μεταβλητή έχει υψηλή συσχέτιση με την τρίτη μεταβλητή (0.9).
 - Η δεύτερη μεταβλητή έχει υψηλή συσχέτιση με την τρίτη μεταβλητή (0.9).
 - Παρόλα αυτά η πρώτη με την δεύτερη μεταβλητή έχουν μηδενική συσχέτιση!
- Εάν χρησιμοποιήσουμε $w=(1,1,-1)$ και ελέγξουμε εάν ισχύει η συνθήκη, μπορούμε να δείξουμε ότι δεν ισχύει, άρα ο πίνακας δεν είναι θετικά ημι-ορισμένος.

Παραγοντικά μοντέλα

- Ο αριθμός των συσχετίσεων που πρέπει να εκτιμηθούν για N μεταβλητές, είναι $N(N+1)/2$.
- Αυτός ο αριθμός μπορεί να είναι πολύ μεγάλος σε κάποιες περιπτώσεις και να δημιουργεί προβλήματα στην εκτίμηση.
- Εναλλακτικά μπορεί να χρησιμοποιηθεί ένα παραγοντικό μοντέλο (factor models).
- Ας υποθέσουμε ότι έχουμε μία σειρά από μεταβλητές (V_i).
- Σε ένα μονο-παραγοντικό μοντέλο, κάθε (V_i) θα είχε ένα κομμάτι το οποίο εξαρτάται από έναν κοινό παράγοντα F και ένα κομμάτι (ε_i) που δεν σχετίζεται με τις υπόλοιπες (V), ούτε με τον κοινό παράγοντα:

$$V_i = b_i F + \varepsilon_i$$

- Τα ε ακολουθούν ανεξάρτητες κανονικές κατανομές.
- Οποιαδήποτε συσχέτιση μεταξύ δύο μεταβλητών (V_i και V_j) προκύπτει από την εξάρτησή τους από τον κοινό παράγοντα και είναι ίση με $b_i b_j V(F)$.

Παραγοντικά μοντέλα (συνέχεια)

- Το βασικό πλεονέκτημα της χρήσης παραγοντικών μοντέλων είναι ότι θέτουμε κάποια δομή στις συσχετίσεις.
- Αυτό βέβαια σημαίνει ότι κάνουμε την υπόθεση ότι η δομή που θέτουμε είναι σωστή.
- Με το μονο-παραγοντικό μοντέλο για N μεταβλητές, αντί να πρέπει να εκτιμήσουμε $N(N+1)/2$ παραμέτρους, θα πρέπει να εκτιμήσουμε $N + 1$: b_1, b_2, \dots, b_N και $V(F)$.
- Τα παραγοντικά μοντέλα μπορούν να επεκταθούν και να συμπεριλάβουν μεγαλύτερο αριθμό παραγόντων (έστω M):

$$V_i = b_{1,i}F_1 + b_{2,i}F_2 + \dots + b_{M,i}F_M + \varepsilon_i$$

- Σε αυτή την περίπτωση, αν οι παράγοντες F_m είναι μεταξύ τους ανεξάρτητοι, η συσχέτιση μεταξύ (V_i και V_j) είναι ίση με:

$$\sum_{m=1}^M b_{m,i}b_{m,j}V(F_m)$$

Από κοινού Κατανομές

- Έστω η τ.μ. X έχει την εξής συνάρτηση πιθανότητας

x	0	1	2
$f(x)$	0.50	0.20	0.30

- Στο παράδειγμα, $f(1) = P(X = 1) = 1/5$.
- Όταν θέλουμε να μελετήσουμε 2 μεταβλητές ταυτόχρονα, χρησιμοποιούμε την από κοινού συνάρτηση πιθανότητας

	x=length			
	129	130	131	
y=width	15	0.12	0.42	0.06
	16	0.08	0.28	0.04

- Στο παραπάνω παράδειγμα, $f(129,15) = P(X = 129, Y = 15) = 0.12$.

Οριακές Κατανομές

- Έστω η από κοινού συνάρτηση πιθανότητας

		x=length		
		129	130	131
y=width	15	0.12	0.42	0.06
	16	0.08	0.28	0.04

- $P(X = 129) = P(X = 129, Y = 15) + P(X = 129, Y = 16) = 0.20$
- Γενικά, η οριακή κατανομή της X

x	129	130	131
$f_X(x)$	0.20	0.70	0.10

- Ομοίως,

y	15	16
$f_Y(y)$	0.60	0.40

Οριακές Κατανομές (2)

- Γενικά, για τον υπολογισμό των οριακών κατανομών (για διακριτές κατανομές),

$$f_X(x) = \sum_y f_{XY}(x, y)$$

and

$$f_Y(y) = \sum_x f_{XY}(x, y)$$

Οριακές Κατανομές (3)

- **Παράδειγμα:** Έστω ότι επιλέγουμε 2 μπαταρίες από ένα σύνολο 12 μπαταριών, όπου 3 είναι καινούριες, 4 χρησιμοποιημένες οι οποίες όμως λειτουργούν και 5 χαλασμένες. Έστω X η τ.μ που συμβολίζει τον αριθμό των καινούριων μπαταριών που επιλέξουμε και Y συμβολίζει τον αριθμό των χρησιμοποιημένων μπαταριών.
- Υπολογίστε την $f_{X,Y}(x, y)$

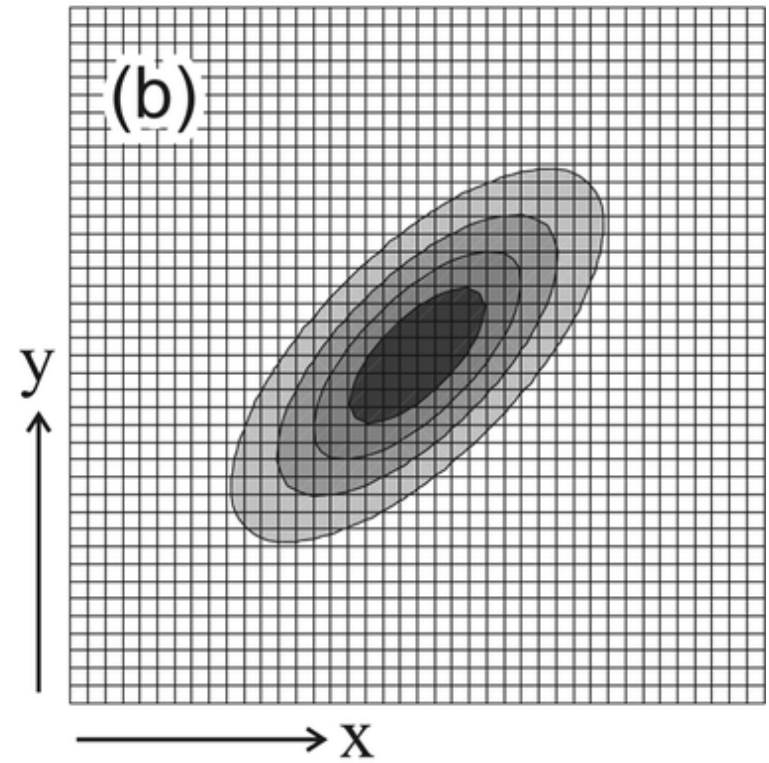
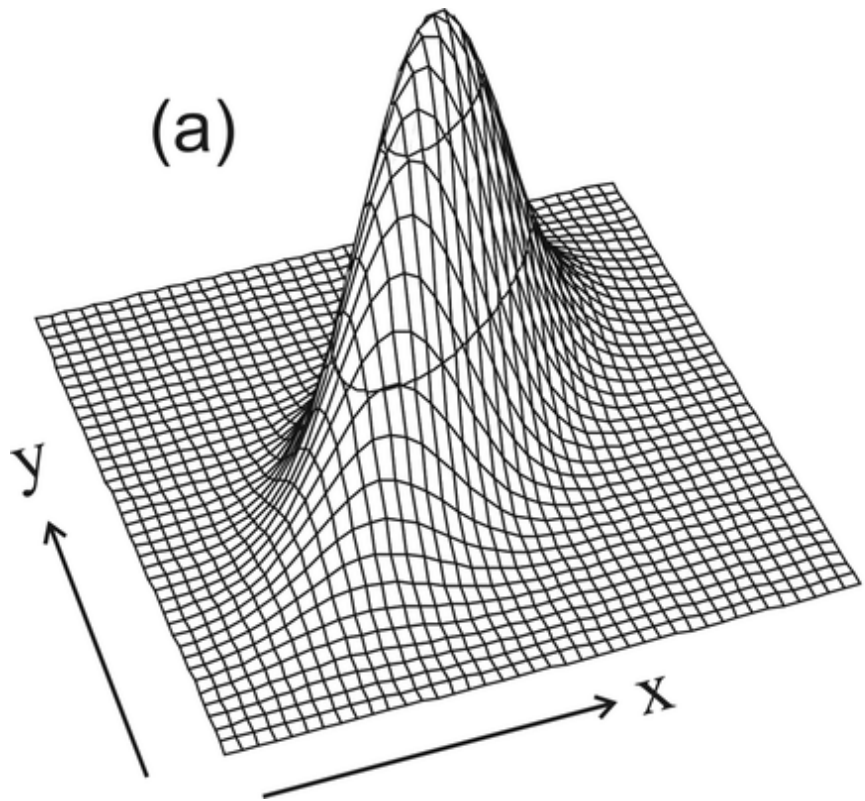
Οριακές Κατανομές (4)

- Η μεταβλητές X και Y παίρνουν τις τιμές 0,1,2. Άρα τα πιθανά ζεύγη είναι τα (0,0), (0,1), (0,2), (1,0), (1,1), (1,2), (2,0), (2,1), (2,2).
- Δεδομένου όμως ότι ο συνολικός αριθμός των επιλεγμένων μπαταριών είναι 2, τα ζεύγη (1,2), (2,1) και (2,2) απορρίπτονται.
- $f(0,1) = P(X = 0, Y = 1) =$ Πιθανότητα στην πρώτη επιλογή να επιλέξω μια από τις 5 ελαττωματικές μπαταρίες και στη δεύτερη να επιλέξω μια από τις 4 ελαττωματικές μπαταρίες $= \frac{5}{12} \frac{4}{11} = \frac{10}{66}$

x= number of *new* chosen

		0	1	2
y=number of <i>used</i> chosen	0	10/66	15/66	3/66
	1	20/66	12/66	
	2	6/66		

Διμεταβλητή Κανονική Κατανομή



Copulas

- Έστω ότι έχουμε χαρτοφυλάκιο με 2 μετοχές.
- Ποια είναι οι πιθανότητες η απόδοση και των δυο μετοχών να μειωθεί κατά 2%?
- Χρειαζόμαστε την από κοινού κατανομή.

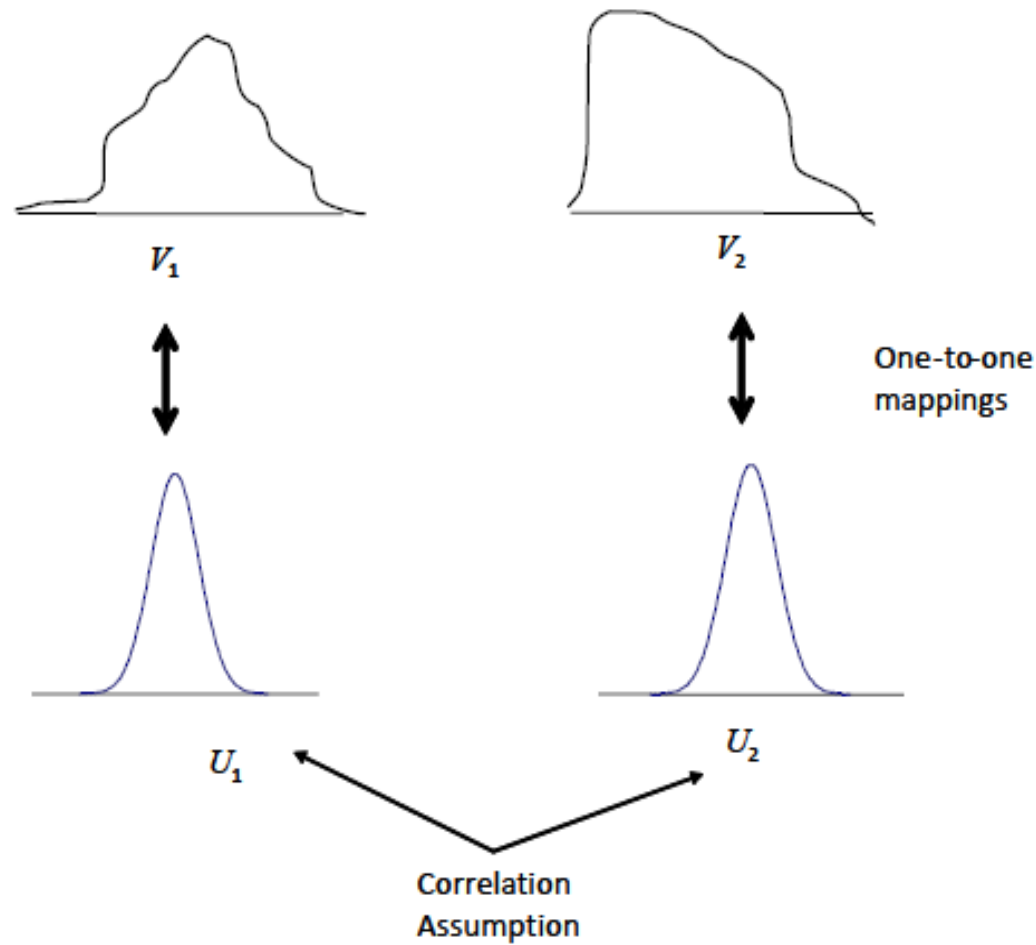
Copulas

- Έστω ότι γνωρίζουμε τις οριακές κατανομές δύο μεταβλητών (έστω V_1 και V_2).
- Πώς θα μπορούσαμε να κάνουμε κάποια υπόθεση για την δομή της συσχέτισης μεταξύ αυτών των 2 μεταβλητών και να ορίσουμε την από κοινού κατανομή τους;
- Μία προφανής επιλογή είναι να υποθέσουμε ότι ακολουθούν μία από κοινού διμεταβλητή κανονική κατανομή (άλλες παρόμοιες κατανομές επίσης)
- Συχνά όμως δεν υπάρχει κάποιος προφανής λογικός τρόπος να ορισθεί μία δομή μεταξύ των συσχετίσεων δύο οριακών κατανομών (εάν για παράδειγμα οι μεταβλητές (V_1 και V_2) δεν κατανέμονται κανονικά).
- Τα Copulas είναι μία τεχνική που μας βοηθάει να ορίσουμε δομή συσχετίσεων (correlation structure) μεταξύ μεταβλητών σε τέτοιες περιπτώσεις.

Copulas: Πώς λειτουργούν

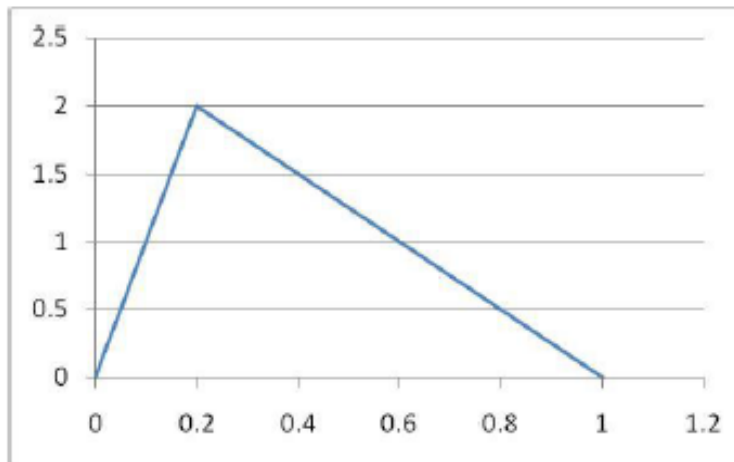
- Για κάθε τιμή των (V_1 και V_2) υπολογίζουμε το αντίστοιχο εκατοστημόριο (percentile) της κατανομής τους.
- Στην συνέχεια, δημιουργούμε δύο νέες μεταβλητές (έστω τις U_1 και U_2) οι οποίες παίρνουν για κάθε ζεύγος τιμής - εκατοστημορίου των αρχικών μεταβλητών, την τιμή της τυποποιημένης κατανομής για το αντίστοιχο εκατοστημόριο.
- Όταν χρησιμοποιείται η τυποποιημένη κανονική κατανομή για την αντιστοίχιση, ονομάζεται ως Gaussian Copula.

Copulas: Πώς λειτουργούν - Διαγραμματικά

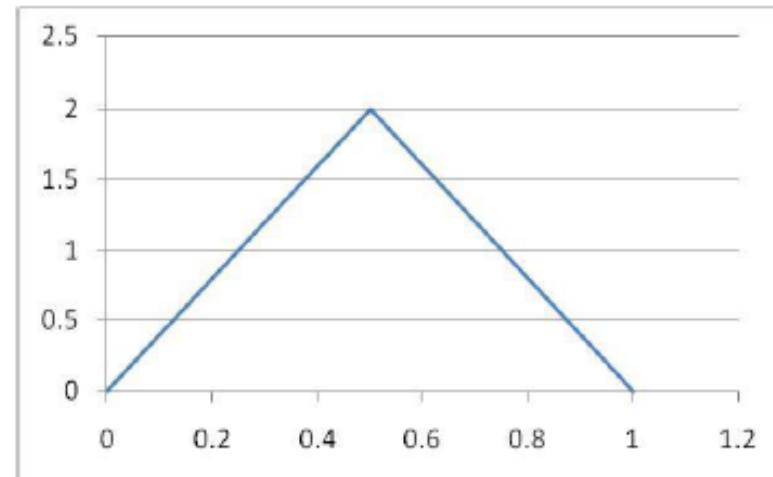


Copulas: Πώς λειτουργούν - Παράδειγμα

- Έστω ότι οι V_1 και V_2 ακολουθούν τις παρακάτω τριγωνικού τύπου κατανομές:



V_1



V_2

Copulas: Πώς λειτουργούν – Παράδειγμα (συνέχεια)

- Η αντιστοίχιση του V_1 σε U_1 :

V_1	Percentile	U_1
0.2	20	-0.84
0.4	55	0.13
0.6	80	0.84
0.8	95	1.64

- Η αντιστοίχιση του V_2 σε U_2 :

V_2	Percentile	U_2
0.2	8	-1.41
0.4	32	-0.47
0.6	68	0.47
0.8	92	1.41

Copulas: Πώς λειτουργούν – Παράδειγμα (συνέχεια)

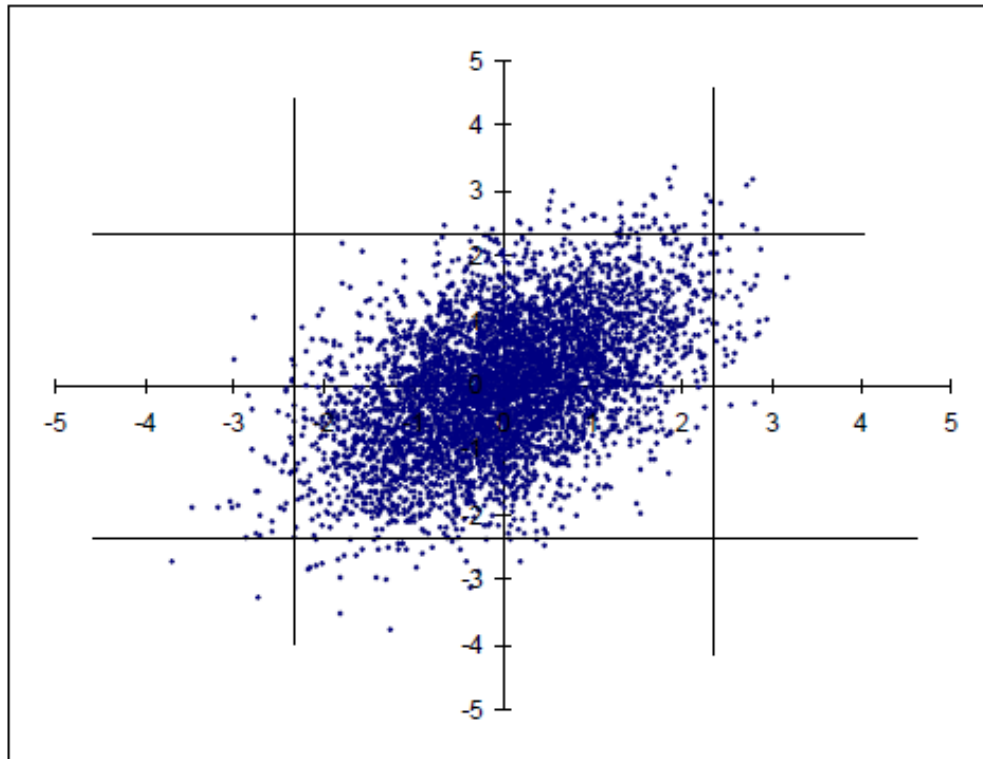
- Ποια είναι η πιθανότητα και οι 2 μεταβλητές να έχουν τιμή < 0.2 ;
- Είναι η πιθανότητα η $U_1 < -0.84$ και η $U_2 < -1.41$.
- Εάν υποθέσουμε ότι η συσχέτιση μεταξύ U_1 και U_2 είναι 0.5, χρησιμοποιώντας την αθροιστική συνάρτηση της διμεταβλητής κανονικής κατανομής:
- $M(-0.84, -1.41, 0.5) = 0.043$ (4.3%).

Copulas: Άλλες Κατανομές

- Θα μπορούσαμε να υποθέσουμε μία οποιαδήποτε άλλη κατανομή για τις U_1 και U_2 .
- Μία εναλλακτική θα ήταν η διμεταβλητή Student t κατανομή.
- Η διμεταβλητή Student t κατανομή έχει μεγαλύτερο tail correlation (συσχέτιση στα άκρα των 2 επίμερους μεταβλητών) σε σχέση με την διμεταβλητή Κανονική κατανομή.
- Σε αυτή την περίπτωση θα έχουμε μία Student t-copula.

Copulas: Άλλες Κατανομές – Διάγραμμα 1

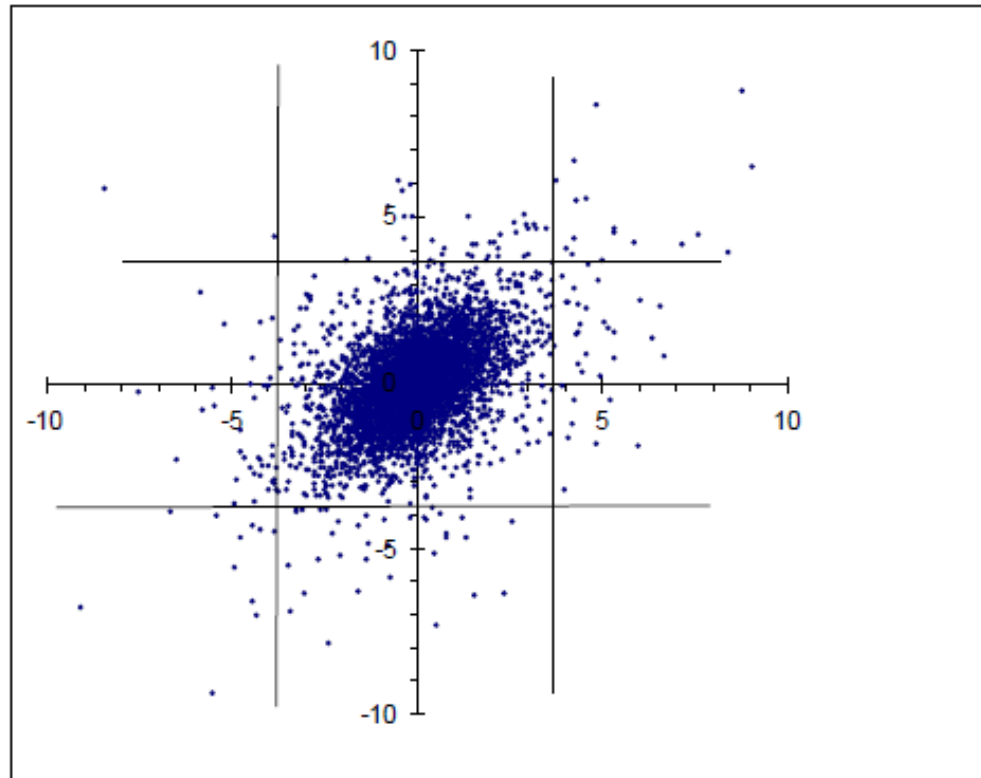
5000 τυχαίες τιμές από μία διμεταβλητή Κανονική κατανομή.



Source: Risk Management and Financial Institutions 3e, Chapter 11, Copyright © John C. Hull 2012

Copulas: Άλλες Κατανομές – Διάγραμμα 1

5000 τυχαίες τιμές από μία διμεταβλητή Student t κατανομή.



Source: Risk Management and Financial Institutions 3e, Chapter 11, Copyright © John C. Hull 2012

Άλλες Copulas

- Θα μπορούσαμε να έχουμε παραπάνω από 2 μεταβλητές.
- Σε αυτή την περίπτωση θα είχαμε με αντίστοιχο τρόπο ένα πολυμεταβλητό Copula.
- Η βασική λειτουργία των copulas είναι στην αντιστοίχιση των τιμών μεταβλητών που δεν ακολουθούν συγκεκριμένες κατανομές, σε τιμές μεταβλητών που ακολουθούν γνωστές κατανομές.