Kovács Edith*

SPECIÁLIS TÖBBVÁLTOZÓS ELOSZLÁSOK MODELLEZÉSE KOPULÁK SEGÍTSÉGÉVEL

Összefoglaló: A cikkben bemutatunk egy módszert az együttes eloszlás modellezésére, azzal a feltételezéssel, hogy ismerjük a peremeloszlás-függvényeket. A többváltozós együttes eloszlásfüggvény teljes mértékben jellemezi a valószínűségi változók összefüggésének mértékét, ezért ez együttes eloszlásfüggvény modellezése az összefüggés mérésének a kiinduló pontja.

Egy portfolió kockázatának a modellezésére idáig leggyakrabban a többváltozós Gauss-eloszlás volt használatos , mivel nagyon sok kényelmes tulajdonsággal rendelkezik. A gyakorlat sajnos bebizonyította, hogy nem mindig illeszthető a valós adatokra.

Egydimenziós valószínűségi változók esetében a normális eloszlással való megközelítés nagyon hasznosnak bizonyul, mivel nagyon sokféle eloszlás közelíthető meg bizonyos feltételek mellett a normálissal.

Mint ismeretes sok olyan *nem* normális együttes eloszlás létezik, amelynek a peremeloszlás- függvényei normálisak. Az egyik fő hiányossága a többváltozós normális eloszlásfüggvénynek, hogy az összefüggés az extremitásokban 0. A valós világban, például a pénzügyekben, sok olyan eset van, amikor a nagyon nagy értékek, illetve nagyon kicsi értékek együttesen fokozott valószínűséggel következnek be. Ezeknek a sajátosságoknak a modellezéséhez fogjuk az archimédeszi kopulák fajtájából a Gumbel, Frank és Clayton családokat használni [4], [5].

Kulcsszavak: normál eloszlás, kopula, extremitásokban való függőség, archimédeszi kopula.

Bevezetés

Dolgozatunk célja két olyan valószínűségi változó együttes eloszlásának a modellezése, amelyek extremitásokban való összefüggést mutatnak (a kiugró értékek egyszerre következnek be) és ezek összefüggésének jellemzése.

A gyakorlatból láthattuk, hogy a kockázat-elemzésben nagyon fontos szerepet játszanak az együtt bekövetkező extrém értékek. Ezeknek figyelembevételével kell majd a modellt megszerkeszteni.

A dolgozatban egy nem normális együttes eloszlás segítségével modellezünk; a fent említett speciális kívánalmakat figyelembe véve, a kétváltozós együttes-eloszlást úgy, hogy a peremeloszlásokat normálisnak tekintjük. Ehhez felhasználjuk az előző cikkünkben [2] bemutatott Gumbel, Clayton és Frank kopulákat.

* *Főiskolai adjunktus, Általános Vállalkozási Főiskola*

A: Rövid emlékeztető a kopulákról

1. A kopulafüggvény értelmezéséről (röviden)

Egyszerűen fogalmazva [3]: több valószínűségi változóhoz tartozó kopulafüggvény az együttes eloszlásfüggvény és a szimpla valószínűségi változók eloszlásfüggvényeit egyesíti magában. Ahogyan azt már [2] bemutattam, a kopulafüggvény segítségével teljesen leírható kettő vagy több valószínűségi változó statisztikai összefüggésének a mértéke.

A kopulafüggvények felhasználásának két fő iránya:

- Meghatározhatjuk a kopulafüggvényt az együttes eloszlásfüggvény ismeretében. (Az együttes el oszlásfüggvény egyértelműen meghatározza a peremeloszlás-függvényeket.
  Ezek ismeretében a Sklar tétel [1] alapján konstruálható egyértelműen a kopulafüggvény.)
- A mintából kapott adatok alapján megválasztjuk a megfelelő kopulát, amelyet egyesítünk a peremeloszlás-függvényekkel. Így modellezhetjük a mintának megfelelő együttes eloszlást.

Definició 1.1: Kopulafüggvénynek nevezünk minden olyan függvényt, amely a következő képpen van értelmezve:

$$C:[0;1]^2 \to [0,1] \text{ és}$$

eleget tesz a következő tulajdonságoknak:

a) $\forall \quad u,v \in I \quad C(u,0)=C(0,v)=0$
b) $\forall \quad u,v \in I \quad C(u,1)=u \qquad C(1,v)=v$
c) $\forall \quad u_1,u_2,v_1,v_2 \in I; \quad u_1 \le u_2,\ v_1 \le v_2$

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

Tétel 1.1. (Sklar) [1]**:** Ha X és Y folytonos valószínűségi változók $H(x,y)$ együttes eloszlásfüggvénnyel és $F(x)$ illetve $G(y)$ peremeloszlás-függvényekkel, akkor létezik egy

$$C:[0;1]^2 \to [0;1]$$

kopulafüggvény, amelyre igaz, hogy

$$\forall\, u,v \in [0;1] \qquad C(u,v)=H\left(F^{-1}(u),G^{-1}(v)\right).$$

Fordítva: Ha $C:[0;1]^2 \to [0;1]$, $C(F(x),G(y))$ egy kopula, akkor létezik egyetlen egy együttes eloszlásfüggvény H(x,y) , F(x) és G(y) peremeloszlásokkal. Erre teljesül:

$$\forall\, x,y \in R \qquad H(x,y)=C(F(x),G(y))$$

2. Az archimédeszi kopulák [5]

Definició 2.1: Azt mondjuk, hogy egy $C:[0,1]^2 \to [0,1]$ kopula archimedeszi, ha létezik egy konvex csökkenő függvény $\phi :]0,1] \to R$ a következő két tulajdonsággal:

$$\phi(1)=0 \quad \text{és} \quad C_\phi(u,v)=\phi^{-1}(\phi(u)+\phi(v))$$

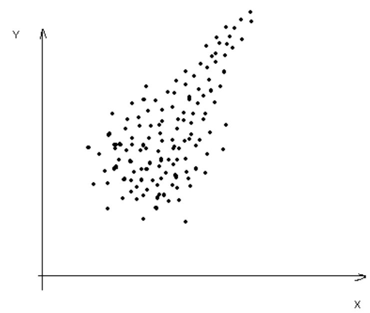
Az archimédeszi kopulafüggvényt egyértelmüen meghatározza a ϕ generálófüggvénye [3], [4].
A továbbiakban a következő archimédeszi kopulacsaládokat fogjuk használni:

1. **Gumbel-család (1960)** $C(u,v)=e^{-\left[(-\ln u)^\alpha+(-\ln v)^\alpha\right]^{\frac{1}{\alpha}}}$ $;\alpha \geq 1$.

$$\phi(v)=(-\ln v)^\alpha$$

Ezt a családot a felső extremitásokban való erős összefüggés jellemezi, mint ez az 1. ábrán látható.

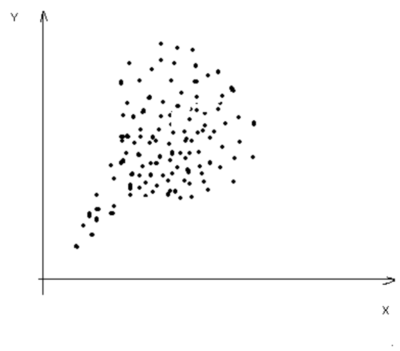
1. ábra



2. **Clayton-család (1978)** $C(u,v)=\left(u^{-\alpha}+v^{-\alpha}-1\right)^{\frac{-1}{\alpha}}; \quad \alpha \geq 0$.

$$\phi(v)=v^{-\alpha}-1 \ .$$

Ezt a családot az összefüggés az alsó extremitásokban jellemzi, ahogyan azt a 2. ábra mutatja.

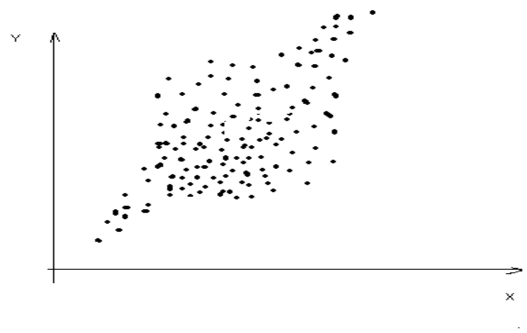


3. Frank család (1978) $C(u,v) = \frac{1}{\alpha}\ln\left(1 + \frac{\left(e^{-\alpha u} - 1\right)\left(e^{-\alpha v} - 1\right)}{e^{-\alpha} - 1}\right) \quad \alpha \in R$

$$\phi(v) = \ln\frac{e^{\alpha u} - 1}{e^{u} - 1}$$

Ez a család szimmetrikusan erős extremitásokban való dependenciát mutat. Ez látható a 3. ábrán.

3. ábra



3. Extremitásokban való dependencia (összefüggés):

- *Felső* extremitásokban való összefüggés, akkor létezik, ha a következő határérték pozitív (vagyis a nagy „kilógó” értékek egyszerre következnek be):

$$\lambda_{upper} = \lim_{\substack{u \to 1 \\ u < 1}} P\left(Y \geq G^{-1}(u) / \ X \geq F^{-1}(u)\right)$$

- *Alsó* extremitásokban való összefüggés akkor van, ha a következő határérték pozitív (vagyis a kicsi „kilógó” értékek egyszerre következnek be):

$$\lambda_{lower} = \lim_{\substack{u \to 0 \\ u > 0}} P\left(Y \leq G^{-1}(u) / \ X \leq F^{-1}(u)\right)$$

Megjegyzések:

- A Gauss-eloszlásnak, amelyhez tartozó korreláció kisebb mint 1 nulla az alsó extremitásokban való dependenciája, úgyszintén a felső extremitásokban való dependenciája.
- Az extremitásokban való összefüggés a kopulán múlik és nem a peremeloszlásokon, ezért a kopulát kell megfelelően megválasztani.

B: Egy speciális együttes eloszlásnak megfelelő kopula modellezése

A továbbiakban abból indulunk ki, hogy X és Y $N(\mu_1, \sigma_1)$ illetve $N(\mu_2, \sigma_2)$ normál eloszlású valószínűségi változók.

Az alapötlet a következő: modellezni fogjuk numerikusan a mintának megfelelő kopulát, szem előtt tartva az extremitásokban való összefüggés típusát. Ennek érdekében a fent bemutatott három kopulacsaládot fogjuk használni.

Tegyük föl, hogy van egy n darab (x_i, y_i) párból álló mintánk. Ezekhez hozzárendeljük a megfelelő u és v értékeket az alábbi képletek szerint:

1.

X	Y	$u = F(x) = \Phi\left(\frac{x-\mu_1}{\sigma_1}\right)$	$v = G(y) = \Phi\left(\frac{y-\mu_2}{\sigma_2}\right)$
x_1	y_1	u_1	v_1
$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	y_n	u_n	v_n

2. Kiszámoljuk az empirikus Kendall-mutatókat:

$$\tau = \frac{c-d}{c+d},$$

ahol c = a mintában lévő konkordáns párok számossága (azok a párok az $\{(x_i, y_i); (x_j, y_j)\}$ párok számossága amelyekre igaz, hogy $(x_i - x_j)\cdot(y_i - y_j) \succ 0$),

d = diszkordáns párok számossága (azok a párok az $\{(x_i, y_i); (x_j, y_j)\}$ párok számossága, amelyekre igaz, hogy $(x_i - x_j)\cdot(y_i - y_j) \prec 0$).

3. Kiszámoljuk az α paramétert, amely a Kendall-féle empirikus mutatónak a függvénye. Kiválasztjuk az adott kopula családból a megfelelő kopulát.

Gumbel	**Clayton**	**Frank**
$\alpha = \frac{1}{1-\tau}; \alpha \in (0, \infty)$	$\alpha = \frac{2\tau}{1-\tau}\quad \alpha \in [0, \infty)$	$\tau = 1 + \frac{4 \cdot D_1(\alpha) - 1}{\alpha}$ where $D_1 = \int_0^\alpha \frac{t}{\alpha(e^t - 1)} dt$ $\alpha \in R$

4. α ismeretében megállapítjuk a generáló függvényt

Gumbel	Clayton	Frank
$\phi(v)=(-\ln v)^{\alpha}$	$\phi(v)=x^{-\alpha}-1$	$\phi(v)=\ln\dfrac{e^{\alpha x}-1}{e^{x}-1}$

5. Felhasználva az első lépésben kiszámított u és v értékeket, most kiszámoljuk a kiválasztott kopula* generáló függvényének a $\phi(u)$ és $\phi(\mathrm{v})$ értékeit:

$\phi(u_1)$	$\phi(v_1)$
$\vdots$	$\vdots$
$\phi(u_n)$	$\phi(v_n)$

6. Az ötödik lépés nyomán kiszámolhatjuk a kopula diszkrét pontokban felvett értékeit.

Gumbel	Clayton	Frank
$C(u_i,v_i)=e^{-\left[(-\ln u_i)^{\alpha}+(-\ln v_i)^{\alpha}\right]^{\frac{1}{\alpha}}}$; $\alpha\geq 1$	$C(u_i,v_i)=\left(u_i^{-\alpha}+v_i^{-\alpha}-1\right)^{\frac{-1}{\alpha}}$; $\alpha\geq 0$	$C(u_i,v_i)=\dfrac{1}{\alpha}\ln\left(1+\dfrac{(e^{-\alpha u_i}-1)(e^{\alpha v_i}-1)}{e^{-\alpha}-1}\right)$ $\alpha\in R$

i=1…n

7. Kiszámoljuk ugyanezeken az $(u_i,\, v_i)$diszkrét pontokon, a függetlenségi kopula által felvett értékeket: $\Pi(u_i,v_i)=u_i.v_i$.

8. Kiszámoljuk ugyanezeken az $(u_i,\, v_i)$diszkrét pontokon, a maximális kopula értékeit $M(u_i,v_i)$, illetve $W(u_i,v_i)$ értékeket.

9. Bevezetünk egy mutatót, amellyel az összefüggés mértékét fogjuk jellemezni.
Ehhez a következő jelöléseket lesznek szükségesek:

- Jelölje $A_i=C(u_i,v_i)-\Pi(u_i,v_i)$.
- Jelölje $\delta_i=\begin{cases}1 & \text{ha } A_i>0\\ 0 & \text{különben}\end{cases}$ és $\overline{\delta}_i=\begin{cases}1 & \text{ha } A_i<0\\ 0 & \text{különben}\end{cases}$

** A Genest, Mac Kay 1993 elmélete és* [8] alapján

Az alapötlet az, hogy számszerűsítsük a modellezett kopula C(u,v) és a függetlenségi kopula közötti eltérést. Ehhez használhatjuk például a az eltérések négyzetének összegét a mintából származtatott pontokon: $\sum_{i=1}^{n}\left(C(u_i,v_i)-\Pi(u_i,v_i)\right)^2$. Minél közelebb áll ez az összeg a nullához, annál közelebb áll a változok kapcsolata a függetlenséghez. A probléma, abból adódhat, hogy ennek a számnak nincsen számszerű felsőhatára.

Tudjuk, hogy minden kopula a minimális és a maximális kopula közé esik, ebből kifolyólag pedig minden (u_i, v_i) pontban számolhatunk egy maximális eltérést:

- Ha a C(u,v) kopula a függetlenségi kopula fölött halad, akkor a maximális eltérést az (u_i, v_i) pontban a $M(u_i,v_i)-\Pi(u_i,v_i)$ adja.
- Ha a kopula a C(u,v) kopula a függetlenségi kopula alatt halad el akkor a maximális eltérés az (u_i, v_i) pontban a $\Pi(u_i,v_i)-W(u_i,v_i)$ adja.

A probléma az, hogy előfordulhat, hogy egy kopula metszheti a függetlenségi kopulát, emiatt pedig megváltoztathatja a függetlenségi kopulához képest a pozícióját. A dolgozat végén található ábrákon bemutatjuk az általunk konstruált kopula és a függetlenségi kopula közötti eltéréseket [a Clayton (CG), Gumbel (CG), Frank (CF)] különböző α értékekre. Ha az eltérések pozitív irányban rajzolódtak ki, akkor az általunk konstruált kopula van fölül (4, 5, 6. ábrák), ha pedig negatív irányban rajzolódnak ki, akkor az általunk konstruált kopula halad el a függetlenségi kopula alatt (7, 8. ábrák), de előfordul, hogy negatív tartományból pozitívba vált át, vagy fordítva. Ilyen esetekben pedig valahol metszi a függetlenségi kopulát, így megváltoztatva a hozzá viszonyított relatív pozícióját.

Ezen sajátosságok figyelembevételével bevezetjük a következő mutatót:

$$\hat{\sigma}_{XY}=\frac{\sum_{i=1}^{n}\left(C(u_i,v_i)-\Pi(u_i,v_i)\right)^2}{\sum_{i=1}^{n}\left[\delta_i\left(M(u_i,v_i)-C(u_i,v_i)\right)^2\right]+\sum_{i=1}^{n}\left[\bar{\delta}_i\left(C(u_i,v_i)-W(u_i,v_i)\right)^2\right]}\ .$$

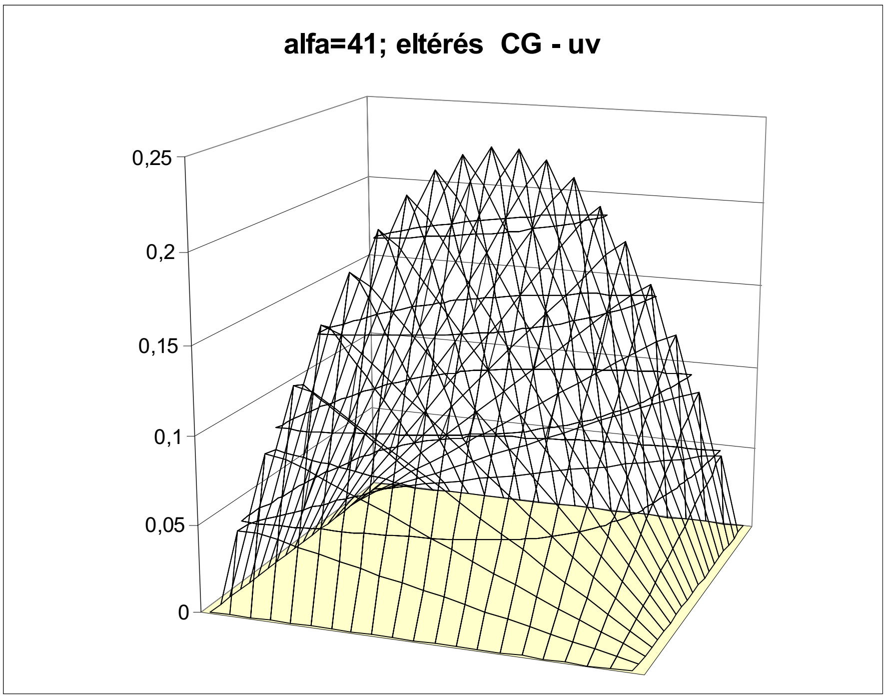
Ez a mutató már nulla és egy közé esik.

C: Következtetések

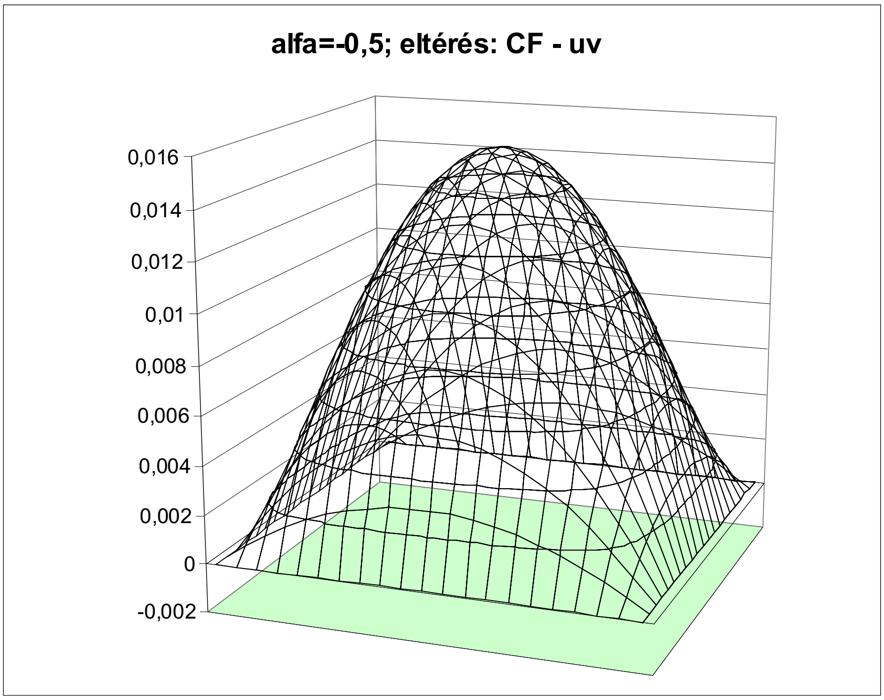
A B: részben bemutatott módszer által numerikusan modelleztünk egy normális peremeloszlású, de nem normális együttes-eloszlásnak megfelelő kopula függvényt, figyelembe véve az extremitásokban megfigyelhető sajátosságokat.

A mintából nyert adatokra támaszkodva bevezetünk egy empirikus mutatót $\hat{\sigma}_{X,Y}$, amely jellemzi az összefüggés mértékét. Minél közelebb áll $\hat{\sigma}_{X,Y}$ az 1-hez, annál erősebb kapcsolatot sejthetünk a két valószínűségi változó között; minél közelebb áll ez a mutató a nullához annál gyengébb kapcsolatot sejtet.

4. ábra * **



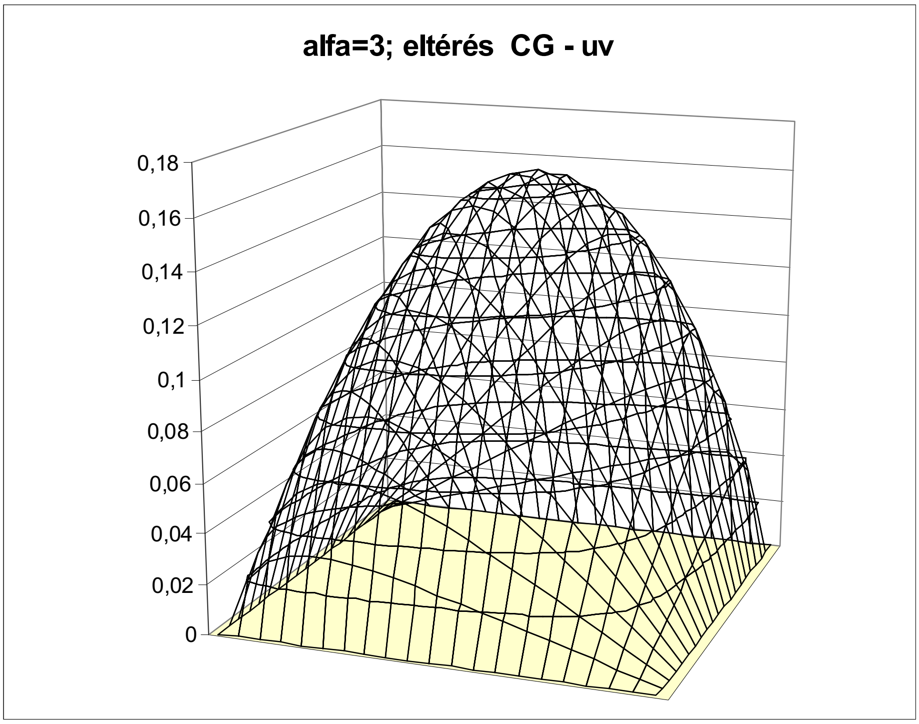
5. ábra * **



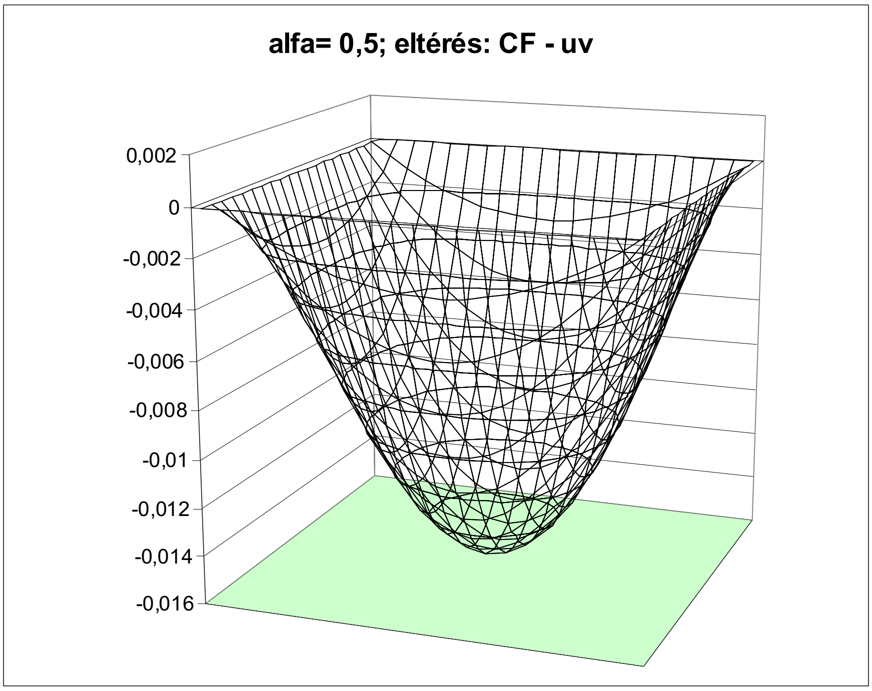
* *A síkban mérjük az u és v értékeket, tehát fekvő négyzet* $[0;1]^2$

** *Érdemes megfigyelni a függőleges tengely értékeit. Minél nagyobbak annál nagyobb az eltérés a függetlenségi kopulától, tehát erősebb az összefüggés. (Megjegyzés: különböző ábrákon különböző a mértékegység.)*

6. ábra * **



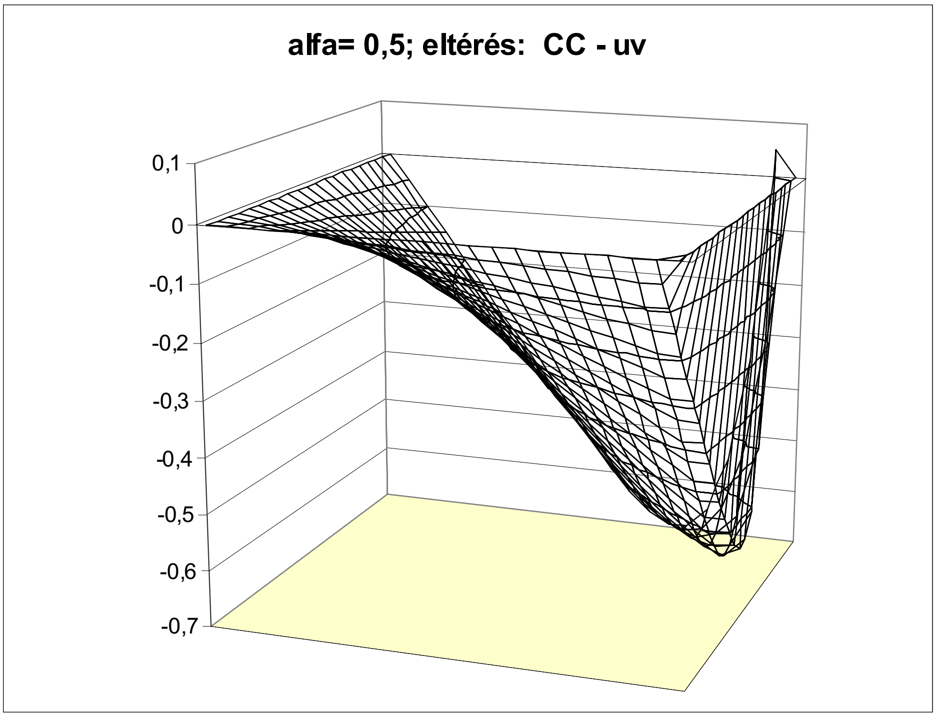
7. ábra * **



* *A síkban mérjük az u és v értékeket, tehát fekvő négyzet* $[0;1]^2$

** *Érdemes megfigyelni a függőleges tengely értékeit. Minél nagyobbak annál nagyobb az eltérés a függetlenségi kopulától, tehát erősebb az összefüggés. (Megjegyzés: különböző ábrákon különböző a mértékegység.)*

8. ábra * **



FELHASZNÁLT IRODALOM:

[1] Sklar, A.: Random variables, distribution functions, and copulas-a personal look backward and forward, in Distributions with Fixed Marginals and Related Topics, ed. By L. Rüschendorff, B. Schweizer, and M. Taylor, pp. 1-14. Institute of Mathematical Statistics, Hayward, CA.

[2] Kovács, E.: Valószínűségi változók együttes eloszlásának illetve összefüggésének jellemzése kopulák segítségével. Általános Vállalkozási Főiskola, Tudományos Közlemények 12. szám, 2005. április.

[3] Genest, C. and Mac Kay: The joy of copulas: Bivariate distributions with uniform marginals, The American Statistician, 40. pp. 280-283. (1986a).

[4] Genest, C. and Mac Kay: Copulae archimédiennes et familles de lois bidimensionell dont les marges sont donnes, The Canadian Journal of Statistics 14. pp. 145-159. (1986b).

[5] Genest, C. and Rivest, L.-P.: Statistical inference procedures for bivariate Archimedian copulas, Journal of the American Statistical Association, 88. pp. 1034-1043.

[6] Embrechts, P. Lindskog, F and Mc Neil A.: Modelling Dependence with Copulas and Applications to Risk Management" Departament of Mathematics ETHZ september 10, 2001. pp. 1-48.

[7] Nelsen, R.: An Introduction to Copulas, Springer, New York, 1998.

[8] Kovács, E.: Algoritmusok az együttes eloszlás modellezéséhez kopulák segítségével. Országos Matematika, Fizika, Infomatika Konferencia. Szeged, 2005.

* *A síkban mérjük az u és v értékeket, tehát fekvő négyzet* $[0;1]^2$

** *Érdemes megfigyelni a függőleges tengely értékeit. Minél nagyobbak annál nagyobb az eltérés a függetlenségi kopulától, tehát erősebb az összefüggés. (Megjegyzés: különböző ábrákon különböző a mértékegység.)*