

Kockázatmodellezés és előrejelzés Pythonnal

Albrecht Martin (HTTKCM)

Szarka Álmos (CFG10)

Hitelek és kockázatok mikro és makro szinten

2022/2023 tavaszi félév

Budapesti Corvinus Egyetem

Tartalom

I.	Bevezetés.....	2
II.	Historikus VaR és a diverzifikáció	2
III.	Hozamszenáriók és a VaR szerepe	4
IV.	EWMA kockázatbecslés.....	6
V.	Kockázatelemzés Machine Learninggel	6
VI.	Melléklet.....	8

I. Bevezetés

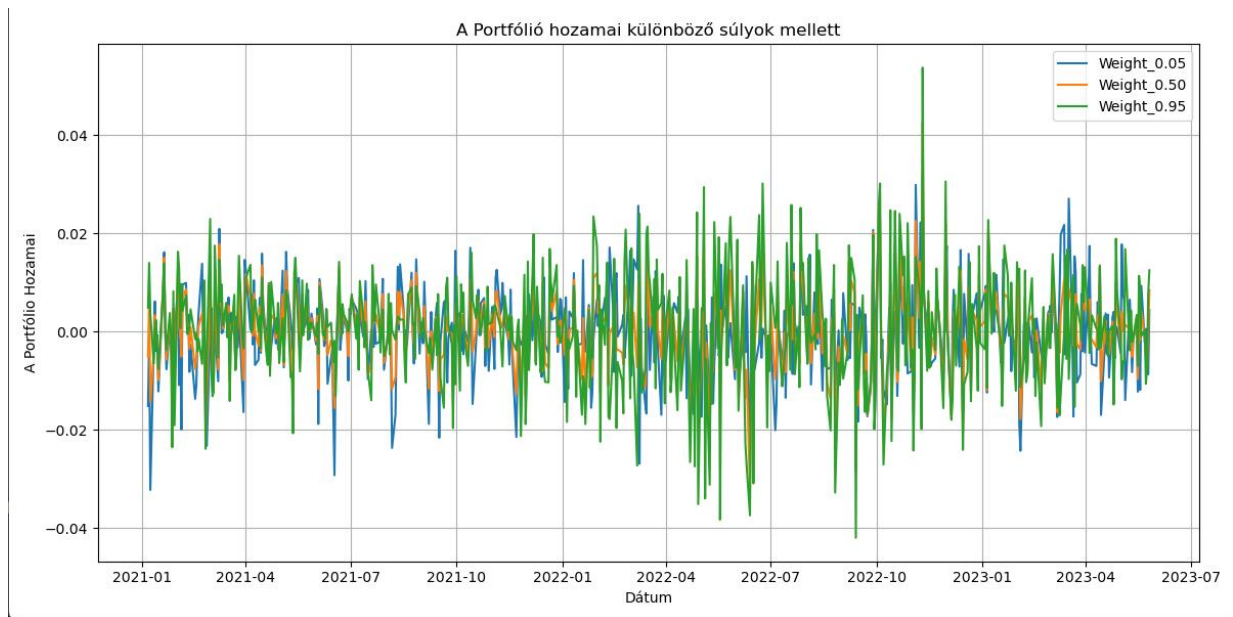
A következőkben kockázatelemzés és előrejelzés céljából fogunk historikus és szimulációs elemzéseket végezni, többek között Value at Risk (VaR), EWMA kockázatbecsléssel és Machine Learning segítségével is. Elemzésünk olyan eszközök kiválasztásával történik, amelyek a köztudatban stabilnak, kevésbé kockázatosnak tekinthetők. Emiatt két ETF-et választottunk terméknek:

- SPY: SPDR S&P 500 ETF Trust, a széleskörben használt, top vállalatokat magába tömörítő S&P 500 indexre épülő ETF
- GLD: SPDR Gold Shares (GLD), az arany árfolyamára épülő ETF

Továbbá célunk volt még az elmúlt évek hatásának kimutatása, így időszaknak egy már a COVID sokk utáni intervallumot választottunk, ami 2021 év elejétől 2023 május végéig tart. Ezt követően két segédfüggvény segítségével ábrázoltuk, hogy a hozamok hogyan alakulnak különböző súlyok mellett.

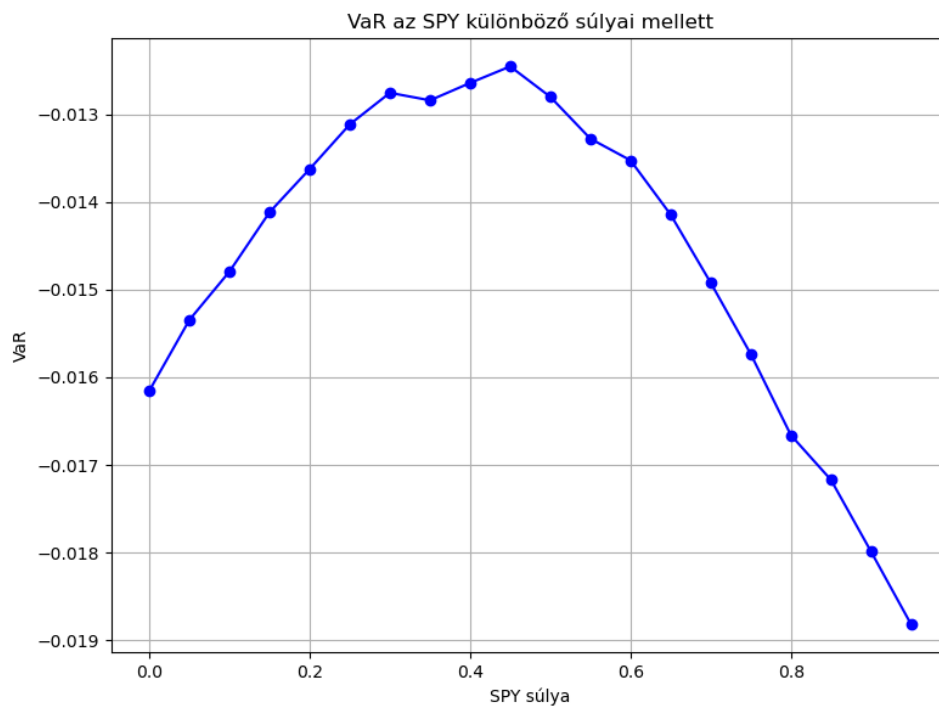
II. Historikus VaR és a diverzifikáció

Első körben a Yahoo Finance oldaláról letöltöttük a megfelelő adatbázisokat, ahonnan a napi záróárakra volt szükségünk. Az adatok beolvasása után egyszerű és loghozamokat is számoltunk a napi árakból, majd egy függvény segítségével különböző súlyok mellett a kialakítandó portfólió hozamát is meghatároztuk, ahol a súly paraméter. Ábráinknál a jobb értelmezés végett csak az egyik súlyt írjuk ki, ami az SPY súlya a portfólióban.



Az ábrán látható, hogy a két papírral néhol jelentősen eltérő hozamokat lehet elérni, valamint az 50%-50% portfólió ezt jól ki tudja egyenlíteni. Kitűnik még, hogy az SPY volatilitása erősebb.

Ezt követően rátérhettünk a historikus VaR-re, melyet több különböző súlykombináció mellett is meghatároztunk. A különböző súlyok melletti portfóliók VaR értékét az alábbi grafikonon ábrázoljuk.



Látható, hogy a VaR értékek $-0,01$ és $-0,02$ között mozognak, 95%-os konfidencia intervallum mellett, melynek a jelentése, hogy nagyjából 5% esély van arra, hogy az adott értéknél nagyobb veszteségünk lesz a befektetéssel. Látható, hogy ahogy az SPY súlya növekszik, úgy egy ideig abszolút értékben csökken a VaR, egészen addig, amíg a súly el nem éri a 45%-50%-ot, majd elkezd ismét negatív irányba menni majdnem elérve a $-0,02$ -öt. Ebből következik, hogy ennél a portfóliónál ezen az időszakon a VaR értékek alapján a legkisebb potenciális veszteség az egyenlő súlyú portfólió környékén van, tehát diverzifikálnunk szükséges. Emellett leolvasható még, hogy az SPY felé jobban eltolt portfólió esetén nagyobb veszteséget érhetünk el, tehát erre az ETF-re jobban érzékeny a portfólióink.

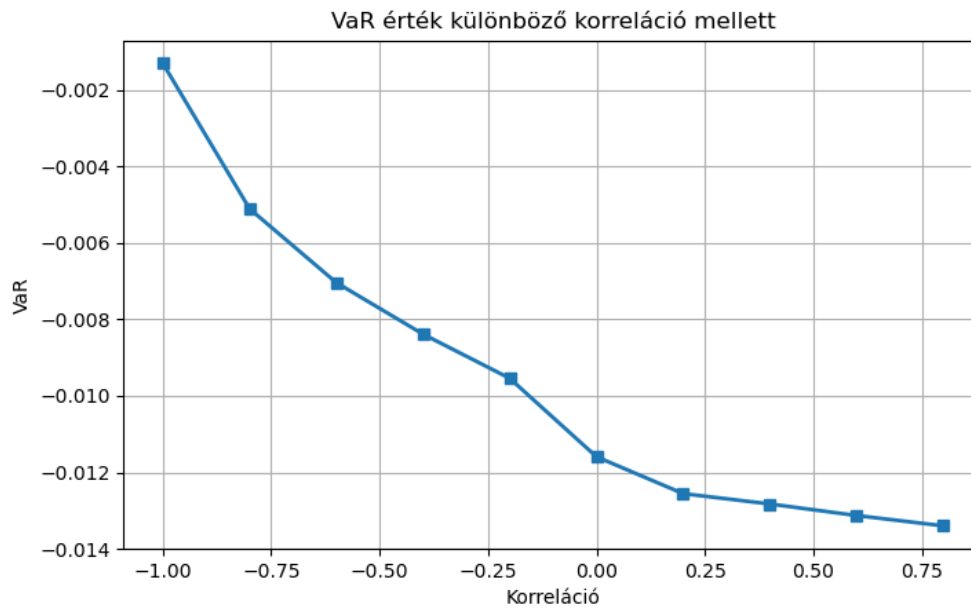
III. Hozamszenáriók és a VaR szerepe

A következő részben egy szimulációt készítünk mellyel az SPY és a GLD jövőbeli hozamait szeretnénk meghatározni a historikus hozamokból, különböző korrelációs feltételezések mellett. Ebben a feladatban rögzítjük a súlyokat, itt is használunk egy feltételezést, majd a súlyokat fordítottan arányossá tesszük a volatilitásokkal. Majd ezen szimulált hozamokra is számolunk VaR-t.

Első körben a hozamokat szimuláltuk, amely függvény első körben egy kovariancia mátrixot számol ki a volatilitásokból és a korrelációból, majd a `numpy.random.multivariate_normal` függvénye segítségével generáljuk le a megfelelő hozamokat, melyek így minták lesznek többváltozós normális eloszlásból a két eszközre, és ezt egy array-be tároljuk. Ehhez a historikus értékek alapján az eszközök várható hozamaira és volatilitásokra van szükség.

Ezt követően a súlyokat 45%-55%-ra fixáltuk, majd a számolt volatilitásokkal fordítottan arányossá tettük őket. Érdekes ábrázolni különböző korrelációs feltételezés mellett a hozamokat. A szélső helyzeteket néztük meg, tehát 1;0 és -1-es korrelációs értékeket. Az ábrák a mellékletben találhatók meg. Látható, hogy 1-es esetben szinte teljesen együtt mozognak csak az SPY esetében nagyobbak a hozamok, -1-es esetben észrevehető az ellentétes mozgás, és a 0-es esetben pedig nincs észrevehető kapcsolat.

Ezek után megírtuk a függvényt, ami most numpy array-ből számolja ki a VaR értéket a megfelelően súlyozott portfólióra különböző korrelációk mellett -1-től egészen 1-ig.

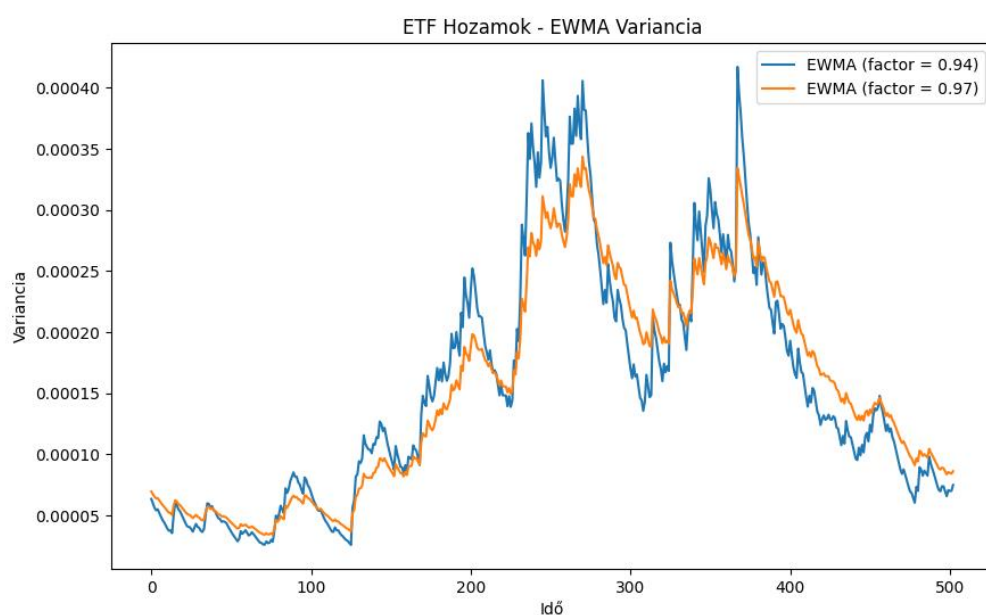


Látható, hogy amikor ellentétesen mozog a két szimuláció, akkor 0 közeli VaR értéket kapunk, hiszen a veszteségek és a nyereségek kioltják egymás hatását. Emellett ahogy növekszik a korreláció úgy nő a VaR is, és a maximális érték pedig a 1-nél lesz, amikor szinte együtt mozog a két termék. Ebből azt a tanulságot vonhatjuk le, hogy diverzifikáció esetén érdemes komolyan venni a korrelációs értékeket, és ha kockázatkerülők vagyunk, akkor csökkenhetjük a kockázatot, ha ellentétesen mozgó papírokat teszünk bele a portfólióba.

IV. EWMA kockázatbecslés

A dolgozat során elemzés alá vetett ETF inntől egyedül S&P500-ra épülő SPY, mely az amerikai tőzsdén jelen lévő 500 legnagyobb piaci kapitalizációval rendelkező vállalatok árfolyamának súlyozott átlagát reprezentálja, ezért gondoltuk alkalmasnak további elemzésre.

Annak érdekében, hogy az idősor elemzése során az adatban jelen lévő trend ne torzítsa az eredményeket, a vizsgált adatok az index napi hozamaiból tevődnek össze. A következő ábra

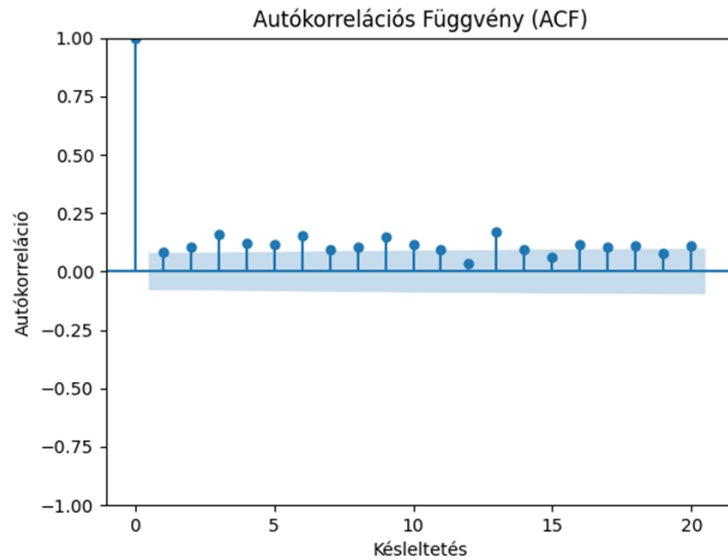


az index hozamain vett exponenciális mozgó átlag értékeit tartalmazza 100 időszaki visszatekintéssel, illetve 0.94 és 0.97-es faktorral.

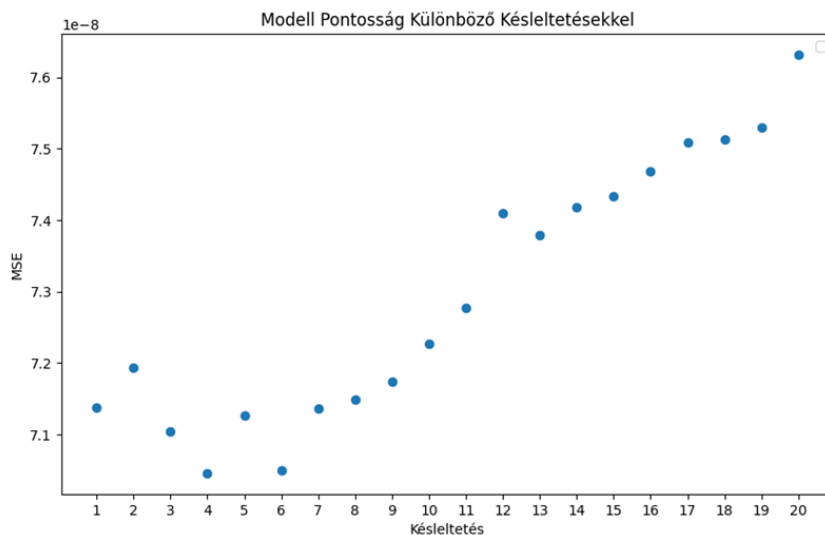
Az ábrán az y tengelyen a mozgóátlag varianciája található, melynek magasabb értékei magasabb volatilitást implikálnak. Emellett jól látható, hogy nagyobb faktor érték lassabb reakciót jelent mind a varianciában fellépő csökkenésre vagy növekedésre. Továbbá egy fontos empirikus tulajdonsága figyelhető meg a pénzügyi idősoroknak, miszerint a volatilitás klasztereződik.

V. Kockázatelemzés Machine Learninggel

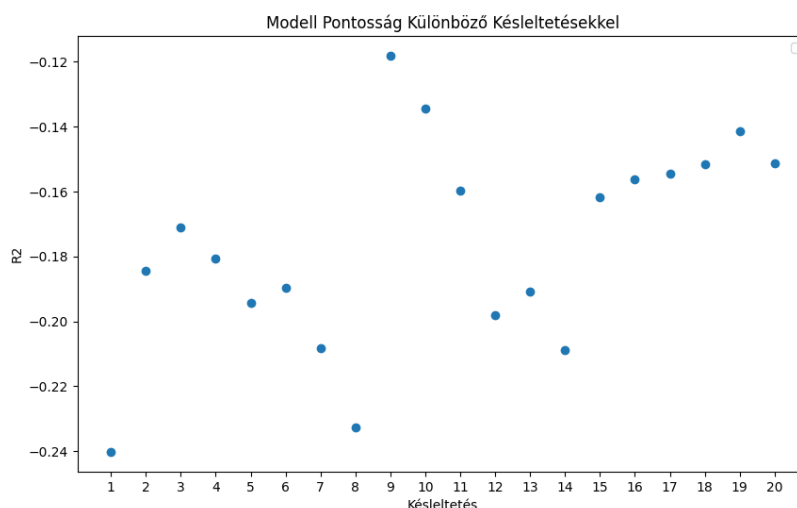
A 4. feladat során az SPY ETF nap hozamainak négyzetei kerülnek vizsgálat alá, melyek a négyzetre emelésnek köszönhetően az adott naphoz tartozó varianciát takarják. A feladat során az említett varianciák modellezése a cél annak késleltetési felhasználásával, ebből adódóan a modellezés előtt érdemes lehet az autokorrelációs függvény vizsgálata, melyet a következő ábra mutat.



Az ábra alapján jól látszik, hogy bár a volatilitás klasztereződik, az autokorreláció egyetlen késleltetés esetében sem szignifikáns, így várhatóan a modell sem fog kimagaslóan jól teljesíteni. A modell tesztelése során 10-es keresztvalidációt használtunk, ahol a teszt halmazon a modell által prediktált értékek és a valódi értékek közötti MSE került kiszámításra, majd átlagolásra. Annak érdekében, hogy az optimális maximális késleltetésszám meghatározható legyen, ezen elemzés minden olyan modellre elkészült, melyek 1-től 20-ig tartalmaznak késleltetéseket.



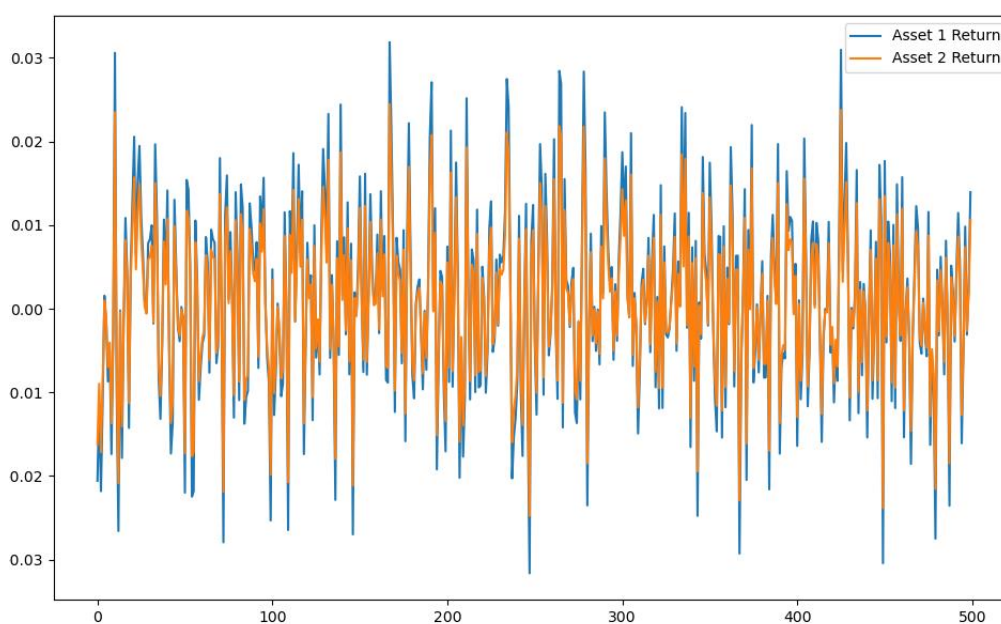
Az ábra alapján látható, hogy a késleltetésszám növelése rontja a modell előre jelzési képességét, és érdemesebb kisebb késleltetésszámot alkalmazni. A grafikon oly módon értelmezhető, hogy egy adott n késleltetésszámmal rendelkező modell minden késleltetést tartalmaz 1-től n -ig. A következő ábra ismét ugyan ezen modelleket tartalmazza, viszont MSE helyett R^2 mértékkel.



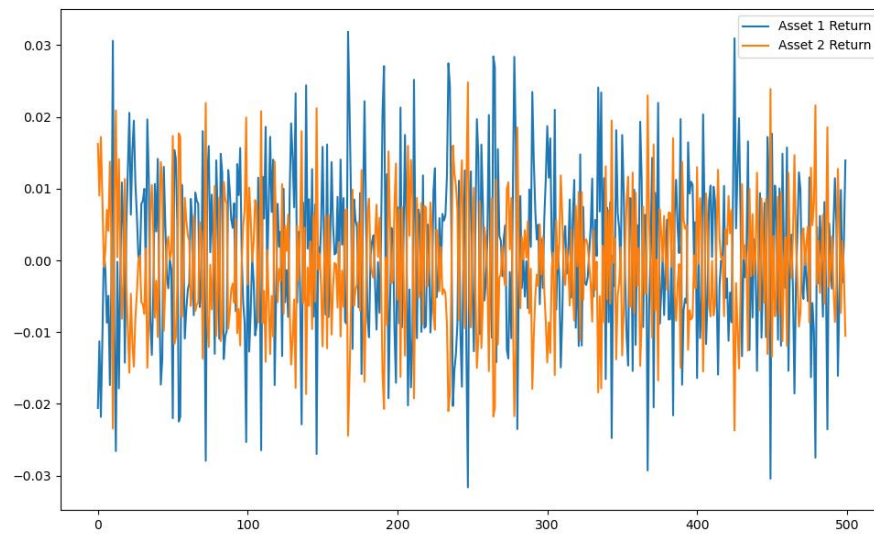
Az utolsó ábra alapján az R^2 kizárólag negatív tartományban található, mely azt jelenti, hogy nem érdemes a jelenlegi független változók használata, az adatok átlaga jobb eredményt hozna, mint a modell becslése.

VI. Melléklet

Szimulációk 1-es korreláció mellett



Szimulációk -1-es korreláció mellett



Szimuláció abban az esetben, ha nincs korreláció

