

NEM-PARAMÉTERES OKSÁG TESZTEK KÉT VÁLTOZÓRA¹

ABALIGETI GALLUSZ
PTE KTK

Tanulmányunkban a klasszikus Granger-féle okság teszt legismertebb nem-paraméteres alternatíváit mutatjuk be, továbbá ismertetünk egy független kopuláción alapuló, önállóan kidolgozott módszert. A tesztek bemutatása után különböző nem-lineáris adatgeneráló folyamatok segítségével szimulált adatsorokon hasonlítjuk össze őket, illetve a referenciaként szolgáló klasszikus okság tesztet. Eredményeink azt mutatják, hogy az új módszer bizonyos esetekben jobban teljesít a korábbiaknál. Végül a Dow Jones index hozama és kereskedési volumene közti oksági viszony vizsgálatán keresztül szemléltetjük, miként lehet gyakorlatban alkalmazni a módszereket. Az empirikus eredmények az irodalomban korábban is kimutatott oksági viszonyok létezését támasztják alá.

1 Bevezetés

A gazdasági modellezés a 20. század második felében a sztochasztikus idősorok felé orientálódott, ami rengeteg új fogalom megjelenését eredményezte. Többek között a sztochasztikus folyamatok és idősorok közti oksági relációk definiálására merült fel igény. Már 1956-ban Wiener is kísérletet tett rá (Wiener, 1956), később Granger egy műhelytanulmányában (Granger, 1963), az áttörést azonban az 1969-es *Econometrica* cikke (Granger, 1969) hozta meg.

Granger eredeti megfontolása gyorsan elterjedt az ökonometriai vizsgálatokban. Ennek oka, hogy – később látni is fogjuk – rendkívül könnyen tesztelhető és nem kell hozzá különösebb számítási kapacitás sem. Ennek következtében gyakorlatilag ez lett az egyetlen okság teszt, ami minden ökonometriai szoftverbe implementálásra került. Az elmúlt 40 évben ugyanakkor megjelentek más alternatívák is okság tesztekre, melyek igyekeznek a Granger-féle okság teszt hiányosságait pótolni, ebben a tanulmányban négy ilyen tesztet mutatunk be és hasonlítjuk össze teljesítményüket a klasszikus Granger-féle okság teszttel.

A precíz elméleti fogalmak előtt próbáljunk intuitíve oksági viszonyt definiálni két idősor között. Arról van tehát szó, hogy tudjuk megragadni az időben eltolt események közötti viszonyt. Erre a legegyszerűbb módszer a

¹Szeretnék köszönetet mondani dr. Rappai Gábornak témavezetőmnek, dr. Kehl Dánielnek kollégámnak valamint a tanulmány lektorának a többszöri alapos átolvasásért és az előremutató megjegyzéseikért. Beérkezett: 2015. május 30. E-mail: xx@ktk.pte.hu.

korreláció pontosabban a parciális korreláltság, hiszen abból már minden más hatás ki van szűrve. A legegyszerűbb, ha egy lineáris modellt írunk fel, ekkor ugyanis a parciális korrelációs együtttható kapcsolatban áll a lineáris modell paramétereivel. Tekintsük tehát a következő modellt:

$$Y_t = \alpha Y_{t-1} + \beta X_{t-1} + \epsilon_t^Y, \quad (1)$$

ahol $\epsilon_t^Y \sim \mathcal{N}(0, 1)$ függetlenek egymástól, korábbi értékeiktől valamint az X_t és Y_t változóktól minden $t = 1, \dots, T$ -re. Ez alapján azt mondhatjuk, hogy X „naivan” *nem oka* az Y -nak, ha $\beta = 0$, hiszen egy adott időpontbeli Y érték és az őt megelőző időpontbeli X érték között nincs – lineáris – kapcsolat.

Általában véve maradunk annál az egyszerű esetenél, amikor két változót vizsgálunk, és az egyik időszori érték a másikhoz képest csupán egy periódussal van késleltetve. Tesszük ezt azért, hogy a lehető legerősebb jelöléseken keresztül tudjuk bemutatni a módszertant. Természetesen minden módszer tetszőlegesen kiterjeszthető a késleltetés hosszának és a változók számának szempontjából, a hivatkozott cikkek ezeket precízen meg is teszik.

A tanulmány szerkezete a következő: a 2. szakaszban a klasszikus Granger-okságról lesz szó röviden, taglalva előnyeit, hátrányait. A 3. szakasz a klasszikus oksági fogalom nem-paraméteres kiterjesztését és a hozzá tartozó teszteket mutatja be, illetve ajánlunk egy, a független kopulán alapuló önállóan kidolgozott módszert, míg a 4. szakaszban Monte-Carlo szimuláció segítségével hasonlítjuk össze az addig bemutatott módszereket különböző adatgeneráló folyamatok esetén, végül az 5. szakaszban egy illusztratív példán mutatjuk be a módszereket.

2 Klasszikus Granger-okság

Az oksági irodalom gyakorlatilag legtöbbször hivatkozott, úttörő cikke Granger (1969) szerint az X idősor oka az Y -nak, ha javítja az előrejelzést az – ebben az esetben – egy periódussal korábbi értékének segítségével. Azaz, ha a múlt információi közül kivesszük az X korábbi értékét, akkor nagyobb lesz a becslésünk varianciája, tehát romlik a becslés:

$$\mathbf{D}^2\epsilon(Y_t | Y_{t-1}, X_{t-1}) < \mathbf{D}^2\epsilon(Y_t | Y_{t-1}), \quad (2)$$

ahol $\mathbf{D}^2\epsilon(Y_t | Y_{t-1}, X_{t-1})$ abból a becslésből származó hibatagok varianciáját jelöli, amiben az Y_t -t magyarázzuk az Y_{t-1} és X_{t-1} változókkal, míg a (2) bal oldalán a restriktív tartalmú modell becslésének varianciája áll. A (2) egyenlőtlenség tagadásaként áll elő a nem-okság definíciója

$$\mathbf{D}^2\epsilon(Y_t | Y_{t-1}, X_{t-1}) = \mathbf{D}^2\epsilon(Y_t | Y_{t-1}), \quad (3)$$

azaz, ha a becslés hibáját nem csökkenti az X_{t-1} változó bevonása (természetesen növelni nem fogja tudni, ezért szerepel egyenlőségjel a két variancia között).

Ez a definíció nem ír elő semmit a modellre, de éppen ezért nem is ad iránymutatást arra nézve, miként kellene tesztelni az okságot. Granger

nyomán tekintsünk ezentúl egyszerűen lineáris esetet, azaz vegyük az (1) modellt, és számoljuk ki először a (3) egyenlet jobb és bal oldalán szereplő varianciák kifejezéséhez szükséges feltételes várható értékeket:

$$\begin{aligned} \mathbf{E}(Y_t | Y_{t-1}, X_{t-1}) &= \mathbf{E}(\alpha Y_{t-1} + \beta X_{t-1} + \epsilon_t^Y | Y_{t-1}, X_{t-1}) = \\ &= \alpha Y_{t-1} + \beta X_{t-1} + \mathbf{E}\epsilon_t^Y = \alpha Y_{t-1} + \beta X_{t-1}, \end{aligned} \quad (4)$$

$$\begin{aligned} \mathbf{E}(Y_t | Y_{t-1}) &= \mathbf{E}(\alpha Y_{t-1} + \beta X_{t-1} + \epsilon_t^Y | Y_{t-1}) = \\ &= \alpha Y_{t-1} + \beta \mathbf{E}(X_{t-1}) + \mathbf{E}\epsilon_t^Y = \alpha Y_{t-1}. \end{aligned} \quad (5)$$

Így a (3) egyenlet alapján, akkor nincs okság X és Y idősorok között, ha felírva a varianciákat

$$\begin{aligned} \mathbf{D}^2\epsilon(Y_t | Y_{t-1}, X_{t-1}) &= \mathbf{D}^2(Y_t - \mathbf{E}(Y_t | Y_{t-1}, X_{t-1})) = \\ &= \mathbf{D}^2(Y_t - \alpha Y_{t-1} - \beta X_{t-1}) = \mathbf{D}^2\epsilon_t^Y, \end{aligned} \quad (6)$$

$$\begin{aligned} \mathbf{D}^2\epsilon(Y_t | Y_{t-1}) &= \mathbf{D}^2(Y_t - \mathbf{E}(Y_t | Y_{t-1})) = \mathbf{D}^2(Y_t - \alpha Y_{t-1}) = \\ &= \mathbf{D}^2(\beta X_{t-1} + \epsilon_t^Y) = \beta^2 \sigma_X^2 + \mathbf{D}^2\epsilon_t^Y, \end{aligned} \quad (7)$$

azok megegyeznek. Tehát az előző szakaszban a nem okságra intuitívan bevezetett kritériumot kaptuk vissza az eredeti Granger-féle definíciót alkalmazva, azaz

$$\mathbf{D}^2\epsilon(Y_t | Y_{t-1}, X_{t-1}) = \mathbf{D}^2\epsilon(Y_t | Y_{t-1}) \Leftrightarrow \beta = 0. \quad (8)$$

Visszatérve a (4) és (5) egyenletekre, rögtön át tudjuk fogalmazni a definíciót, hiszen

$$\mathbf{E}(Y_t | Y_{t-1}, X_{t-1}) = \mathbf{E}(Y_t | Y_{t-1}) \Leftrightarrow \beta = 0 \quad (9)$$

feltétel is igaz lesz. Vagyis, ha nem változik a feltételes várható értéke Y idősornak az X egy periódussal korábbi értékének ismeretében, akkor X nem oka Y -nak, vagy másképpen ha X nem módosítja Y várható értékét, akkor nem is oka annak.

Végül egy harmadik módon is ki lehet fejezni a nem-okságot, mégpedig a korábban már említett parciális korreláció segítségével:

$$\begin{aligned} \mathbf{E}(Y_t X_{t-1} | Y_{t-1}) &= \mathbf{E}((\alpha Y_{t-1} + \beta X_{t-1} + \epsilon_t^Y) X_{t-1} | Y_{t-1}) = \\ &= \mathbf{E}(\alpha Y_{t-1} X_{t-1} + \beta X_{t-1}^2 + \epsilon_t^Y X_{t-1} | Y_{t-1}) = \\ &= \alpha Y_{t-1} \mathbf{E}(X_{t-1}) + \beta \mathbf{E}(X_{t-1}^2) + \mathbf{E}(\epsilon_t^Y) \mathbf{E}(X_{t-1}) = \beta \sigma_X^2, \end{aligned} \quad (10)$$

amiből a nem okság feltétele:

$$Y_t \perp X_{t-1} | Y_{t-1} \Leftrightarrow \beta = 0. \quad (11)$$

Ezek alapján már tudunk adni egy definíciót a klasszikus nem-okság fogalmára, ami ebben az egyszerű 2-változós, egy periódus késleltetést tartalmazó esetben az alábbi:

1. Definíció. Az X idősor klasszikus vagy gyenge értelemben nem-oka az Y -nak, ha az alábbi ekvivalens tulajdonságok valamelyike teljesül²

- (i) $\mathbf{D}^2\epsilon(Y_t | Y_{t-1}, X_{t-1}) = \mathbf{D}^2\epsilon(Y_t | Y_{t-1})$,
- (ii) $\mathbf{E}(Y_t | Y_{t-1}, X_{t-1}) = \mathbf{E}(Y_t | Y_{t-1})$,
- (iii) $Y_t \perp X_{t-1} | Y_{t-1}$.

Gyakorlatban a definíció teljesülését – egyszerű esetben a $\beta = 0$ körülmény fennállását – az (i) tulajdonság vizsgálatára vezetjük vissza. F-próbával tesztljük az X_{t-1} bevonásának szignifikáns varianciacsökkentő hatását (Wald-teszt).

A klasszikus oksági fogalmat lezárandó vegyük sorba, milyen előnyökkel, hátrányokkal bír ez a megközelítés. Mellette szól, hogy egyszerű, könnyen érthető fogalomról van szó, a hozzá tartozó tesztek egyszerűen elvégezhetőek ráadásul a legtöbb statisztikai, ökonometriai szoftver támogatja őket, ráadásul a teszt a mintaelemszámra nézve robusztus. Ugyanakkor a linearitás feltételezése nagyon megkötí a kutató kezét, miközben egyáltalán nem biztos, hogy modellezni is kívánja a folyamatokat, csupán kideríteni változók közötti okság meglétét. További hátránya, hogy magasabb rendű momentumokban bekövetkező hatásokat nem mutatja ki. Napjainkban a pénzügyi idősorok kapcsán már nemcsak a várható érték (vagy valamilyen származtatott változó várható értéke) fontos, ha csak a VaR-ra gondolunk, máris világos, hogy a szórás központi szereppel bír a modellekben, de hasonlóan lényeges lehet a döntéshozóknak a csúcossága vagy ferdesége egy-egy hozamidősrnek. Az ezeket befolyásoló hatások kimutatására sajnos a klasszikus Granger-féle teszt – közvetlenül – nem alkalmas. Erre kínál megoldást a következő szakaszban bemutatandó nem-paraméteres változata a Granger-okságnak.

3 Nem-paraméteres Granger-okság

Az előző szakasz végén láttuk, hogy a klasszikus Granger-okság – a sok jó tulajdonsága mellett – rengeteg hiányossággal bír, ezeket orvosolandó merült fel az erős értelemben vett oksági fogalom bevezetése.

Az 1. Definícióból kiderült, hogy a gyenge okságot jól lehet jellemezni a feltételes várható érték és a feltételes kovariancia fogalmával. Ez utóbbi fogalmak szigorítása merült fel Granger–Newbold (1977) könyvben, ahol a szerzőpáros definiálja az erős okságot:

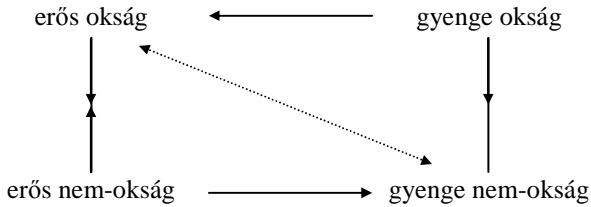
2. Definíció. Az X erősen nem oka az Y -nak, ha az alábbi két ekvivalens tulajdonság valamelyike teljesül minden $t = 1, \dots, T-r$ ³

- (i) $\mathcal{L}(Y_t | Y_{t-1}) = \mathcal{L}(Y_t | Y_{t-1}, X_{t-1})$,
- (ii) $Y_t \perp\!\!\!\perp X_{t-1} | Y_{t-1}$.

²Az (ii) átírást lásd Granger (1963) cikkben, az (iii) átírást lásd Florens–Mouchart (1985) cikkben.

³Természetesen az (i) és (ii) egyenletek ekvivalenciája könnyen belátható, azért emeltük ki mindkét megfogalmazási módot, mert az irodalomban mindkettőt gyakran használják.

A definíció első állítása a feltételes várható értéket szigorította feltételes eloszlássá, míg a második a feltételes korrelálatlanságot feltételes függetlenséggé. Rögtön látszik, hogy ha X erősen nem-oka Y -nak, akkor gyengén sem oka, hiszen ha két eloszlás megegyezik, akkor azok várható értékei is megegyeznek. Másrészt a ha X gyengén oka Y -nak, akkor erősen is oka hiszen, ha a feltételes várható értékek nem egyenlőek akkor a feltételes eloszlások sem lehetnek egyenlőek. Ezeket a tulajdonságokat foglaltuk össze az alábbi ábrán.



Ugyanakkor megfordítva ezek az állítások nem igazak, példaként tekintsük a következő idősorokat:

$$Y_t = X_{t-1} \cdot \epsilon_t^Y$$

$$X_t = \epsilon_t^X ,$$

ahol $\epsilon_t^X, \epsilon_t^Y \sim \mathcal{N}(0, 1)$ korrelálatlan fehér zajok. Vizsgáljuk először is a gyenge okságot! Az 1. definíció (ii) pontjában szereplő egyenlőség két oldalát kifejtve:

$$\mathbf{E}(Y_t | Y_{t-1}, X_{t-1}) = \mathbf{E}(X_{t-1} \cdot \epsilon_t^Y | Y_{t-1}, X_{t-1}) = \mathbf{E}\epsilon_t^Y \cdot X_{t-1} = 0 \quad (12)$$

$$\mathbf{E}(Y_t | Y_{t-1}) = \mathbf{E}(X_{t-1} \cdot \epsilon_t^Y | Y_{t-1}) = \mathbf{E}(\epsilon_{t-1}^X \cdot \epsilon_t^Y) = \mathbf{E}\epsilon_{t-1}^X \cdot \mathbf{E}\epsilon_t^Y = 0 , \quad (13)$$

mivel a két feltételes várható érték megegyezik, a definíció értelmében az X gyengén nem oka az Y -nak. Nézzük most meg az erős okságot a 2. definíció (i) pontja alapján:

$$\mathcal{L}(Y_t | Y_{t-1}, X_{t-1}) = \mathcal{N}(0, X_{t-1}^2) \quad (14)$$

$$\mathcal{L}(Y_t | Y_{t-1}) = \mathcal{L}(Y_t) = \mathcal{L}(\epsilon_t^Y) \cdot \mathcal{L}(\epsilon_{t-1}^X) = \mathcal{N}(0, 1) \cdot \mathcal{N}(0, 1) , \quad (15)$$

utóbbi viszont még csak nem is normális eloszlás, tehát erős értelemben oka X az Y -nak.

A gyenge vagy klasszikus értelemben vett oksági definíciókhoz lineáris modellen keresztül jutottunk el, míg az erős okság definíciójához nem kellett semilyen modellfeltevéssel élni, ezért a továbbiakban az erős okságra a szakirodalomban elterjedt *nem-paraméteres okság* névvel hivatkozunk.

A 2. definíció (ii) pontja alapján, egyszerűen ki lehet mondani a mintabeli nem-paraméteres nem-okságra vonatkozó nullhipotézist:

$$H_0 : \mathcal{L}(Y_t | Y_{t-1}, X_{t-1}) = \mathcal{L}(Y_t | Y_{t-1}) . \quad (16)$$

Sajnos ennek az ellenőrzése rendkívül nehézkes, éppen ezért rengeteg, egymástól akár jelentősen is különböző módszer is ismert az irodalomban sokszor egészen mély matematikai ismeretek igényelve. Közös ugyanakkor a tesztekben, hogy – éppen a nem-paraméteres – mivoltuk miatt csupán valamilyen határeloszlást tudnak követni a tesztstatisztikák, amik pedig viszonylag nagy mintaelemszámot követelnek meg, ez tovább szűkíti ezen tesztek alkalmazhatóságát. Utóbbi problémát súlyosbítja, hogy minél több késleltetést szerepeltetünk a vizsgálatban, a felhasználható része a mintának is annál rövidebb lesz.

Röviden áttekintve a módszerek Su–White (2007) cikkben a szerzők a feltételes eloszlások egyezőségét a változókhoz tartozó feltételes karakterisztikus függvények eltérésén keresztül igyekeznek vizsgálni, míg Sun (2008) funkcionálanalízisbeli ismeretekre építve a kovariancia operátor segítőjével oldja meg a nem-linearitás problémáját. Van azonban egy jobban körülhatárolható halmaza a módszereknek, amikor a sűrűségfüggvények viszonyait vizsgálják a szerzők, utóbbiak információtartalmát sűrítve, átalakítva kezelik a problémát. Jelen tanulmány kereteibe ez utóbbi módszerek összehasonlítása fért bele.

A szakasz végén vegyük sorba, milyen előnyökkel, illetve hátrányokkal bír a nem-paraméteres okság fogalom a klasszikus értelemben vett Granger-oksággal szemben. A fogalom előnyei:

- Semmilyen feltevással nem kell élni arra vonatkozólag, hogy milyen modell illeszkedik az időszori értékekre, illetve milyen eloszlást követnek, ez adja a nem-paraméteres jellegét a tesztnek.
- Míg a klasszikus Granger-teszt csupán az „okozati” idősor várható értékében bekövetkező változást jelzi, azaz csak az első momentumban történő okságot mutatja ki, addig a nem-paraméteres okság tesztek a magasabb rendű momentumokban is jeleznek. Ez elsősorban a rövid távú, magas frekvenciás – általában pénzügyi – adatokban rendkívül fontos, ahol a várható érték olykor teljesen irreleváns, sokkal fontosabb a szórás.

Természetesen hátrányai is vannak ennek a megközelítésnek:

- Látni fogjuk a következő szakaszokban, hogy ezek a próbák összetett számításokat igényelnek, amihez társul a szoftveres támogatás hiánya is, így egy gyakorlati felhasználónak nehéz dolga van, ha alkalmazni szeretné őket.
- Ebben az esetben két eloszlást hasonlítunk össze, tehát az idősoroktól elvárt az erős stacionaritás ahhoz, hogy a próbák megbízható eredményt

adjanak⁴. Szemben a klasszikus Granger-féle teszttel, amihez csupán a gyenge stacionaritás szükséges.

- Kis minták esetében erősen megkérdőjelezhetőek az eredmények, amiket a nem-paraméteres tesztek szolgáltatnak.

A következőkben két, merőben más megközelítést mutatunk be a nem-paraméteres okság tesztelésére. Az egyes megközelítésekben gyakorlatilag az a különbség, hogyan tudják az eloszlásokban rejtőző információt úgy sűríteni, hogy végül tesztelhető legyen az eredeti nullhipotézis – azaz a feltételes eloszlások egyezősége. Az első információsűritő módszer a korrelációs integrál, melynek kapcsán két, egymással közeli kapcsolatban álló tesztet mutatunk be. Míg a második információsűritésre alkalmazott módszer a kopulák segítségével történik, gyakorlatilag arról van szó, hogy a kopulák segítségével nem véges tartójú sűrűségfüggvények „beszoríthatóak” az egységnégyzetbe (kockába), s így az eloszlások vizsgálata leegyszerűsödik.

3.1 Korrelációs integrál alapú tesztek

A (16) egyenletben megfogalmazott nullhipotézis könnyebb kezelhetősége érdekében tegyük fel, hogy az Y_t és X_t változóknak létezik sűrűségfüggvénye. Így a feltételes függetlenség akkor teljesül, ha

$$f_{Y_t|Y_{t-1}, X_{t-1}}(y_t | y_{t-1}, x_{t-1}) = f_{Y_t|Y_{t-1}}(y_t | y_{t-1}), \quad (17)$$

ahol y_t, y_{t-1}, x_{t-1} tetszőleges valós számok. Azaz az X változó nem oka az Y -nak, ha az iménti sűrűségfüggvények minden pontban megegyeznek. Ezzel kapcsolatban – többek között – két nagy probléma merül fel:

- Empirikus esetben egy folytonos változó a legritkább esetben veszi fel kétszer ugyanazt az értéket, így a mintában annyi különböző feltételünk lesz a sűrűségfüggvényekben, amennyi a mintaelemszáma, ebből következően egy elemű „almintákon” kellene tesztelni a függetlenséget – ugyebár minden különböző feltételre –, ami értelmetlen feladat.
- Ha az előbbi problémát megoldottuk, akkor is a sűrűségfüggvények teljes tartóján kellene összehasonlítani az egyenlet jobb és bal oldalát, ami szintén nagyon nehéz.

Ez adja a motivációt, egy olyan fogalom bevezetéséhez, ami egyrészt a feltételben „megengedőbb” a szigorú egyenlőségnél, másrészt van annyi információtartalma, hogy ki lehessen váltani vele a sűrűségfüggvényt, ugyanakkor ne egy függvény legyen – amit ismét pontonként kellene összehasonlítani –, hanem egy, a mintát jellemző skálár.

Ezeknek az elvárásoknak próbál megfelelni a korrelációs integrál, melyet Baek-Brock (1991) cikk alapján a következőképpen definiálunk:

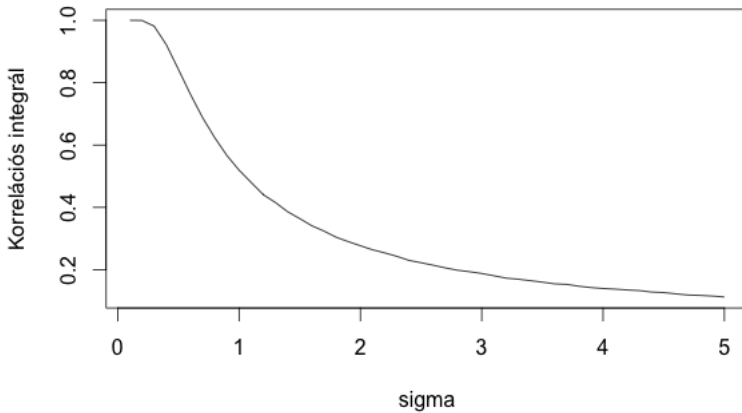
⁴Mindhárom bemutatásra kerülő módszer esetében elvárás, lásd (Hiemstra–Jones, 1994), (Diks–Panchenko, 2005), (Taamoutia et al., 2014).

3. Definíció. Az X d -dimenziós valószínűségi változóhoz tartozó korrelációs integrál a következő

$$C_X(\delta) = \mathbf{P}\left(\max_{i=1}^d |X_i^1 - X_i^2| < \delta\right), \quad (18)$$

ahol $X^1, X^2 \sim X$ független valószínűségi változók és $\delta > 0$.

Természetesen $d = 1$ esetben a definíció egyszerűen $C_X(\delta) = \mathbf{P}(|X^1 - X^2| < \delta)$. Ez gyakorlatilag egy, az eloszlásra jellemző mérőszám (mint pl. az eloszlásnak egy momentuma), azt fejezi ki, mennyire koncentrálnak az adott eloszlásban az értékek. Vegyük példaképpen a normális eloszlást, ekkor a szórás függvényében kifejezhető a korrelációs integrál, amit az 1. ábra szemléltet.



1. ábra Normális eloszláshoz tartozó korrelációs integrál különböző szórások mellett ($\delta = 1$, $n = 10000$)

Jól látható, hogy alacsony szórás mellett a korrelációs integrál értéke egyhez közeli, ahogy nő a szórás, úgy kevésbé koncentráltak az értékek, és ezzel párhuzamosan csökken a korrelációs integrál értéke is.

Lássuk, miként kapcsolódik a korrelációs integrál fogalma eloszlások egyenlőségének teszteléséhez, eleget tud-e tenni elvárásainknak. Legyen $f_X(x)$ az X valószínűségi változóhoz tartozó sűrűségfüggvény, először is vezessük be a következő indikátorfüggvényt, amit a későbbiekben is használni fogunk:

$$\mathbf{1}_\delta(x, y) = \begin{cases} 1, & \text{ha } |x - y| < \delta \\ 0, & \text{egyébként.} \end{cases} \quad (19)$$

Vegyünk egy n elemű mintát X -ből, jelölje x_i az i -edik mintaelemet, és becsljük meg minden mintaelemre a hozzá tartozó sűrűségfüggvény értéket:

$$f_X(x_i) \approx \frac{1}{2\delta} \frac{1}{n-1} \sum_{j \neq i} \mathbf{1}_\delta(x_i, x_j) = \hat{f}_X(x_i). \quad (20)$$

Most átlagoljuk le minden mintaelemre a hozzá tartozó becslt sűrűségfüggvény értékét:

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n f_X(x_i) &\approx \frac{1}{n} \sum_{i=1}^n \hat{f}_X(x_i) = \\ &= \frac{1}{2\delta} \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j \neq i} \mathbf{1}_\delta(x_i, x_j) = \hat{C}_X(\delta) \approx C_X(\delta), \end{aligned} \tag{21}$$

ezzel végérvényben egy becslőfüggvényét kaptuk a korrelációs integrálnak. A (21) egyenletből az is látszik, hogy ha két eloszlás megegyezik, akkor nyilván a hozzájuk tartozó korrelációs integrál is, tehát ha a korrelációs integrálra vonatkozó nullhipotézist vetünk el, akkor az eredeti eloszlásokra is el kell vetni. Ennek az állításnak a megfordításáról az irodalomban nem találtunk eredményt. Úgy sejtjük, általában nem igaz, de „hétköznapi” eloszlásokra valamint elég kicsi δ mellett bizonyos mértékben igaz a megfordítás is, hiszen egyébként nem lenne erre az erre épülő teszteknek.

Gyakorlatilag ez a logikája mindkettő korrelációs integrálra építő tesztnek, amiket a következő szakaszokban bemutatunk.

3.1.1 Hiemstra–Jones-teszt

A továbbiakhoz alakítsuk át a (17) egyenletet a következő – tesztelendő – formára:

$$H_0 : f_{Y_t, X_{t-1} | Y_{t-1}}(y_t, x_{t-1} | y_{t-1}) = f_{Y_t | Y_{t-1}}(y_t | y_{t-1}) f_{X_{t-1} | Y_{t-1}}(x_{t-1} | y_{t-1}), \tag{22}$$

ez a valószínűségi változók nyelvén pontosan azt jelenti, hogy az Y_t az X_{t-1} -től független Y_{t-1} ismeretében.

A Hiemstra–Jones-teszt (Hiemstra–Jones, 1994) ebből az alakból indul ki, és maga a módszer a következő észrevételen alapul⁵:

1. Lemma. *Amennyiben a (22) igaz, úgy különböző t és s időpontokra*

$$\begin{aligned} &\mathbf{P}(|Y_t - Y_s| < \delta, |X_{t-1} - X_{s-1}| < \delta \mid Y_{t-1} = Y_{s-1} = y) = \\ &= \mathbf{P}(|Y_t - Y_s| < \delta \mid Y_{t-1} = Y_{s-1} = y) \mathbf{P}(|X_{t-1} - X_{s-1}| < \delta \mid Y_{t-1} = Y_{s-1} = y). \end{aligned} \tag{23}$$

Bizonyítás. Az F1 függelékben közöljük. □

A szakasz elején bevezetett korrelációs integrál segítségével a (23) egyenlőség, mint nullhipotézis már „majdnem” tesztelhető. A Szerzők ötlete az, hogy ezt a feltételben szereplő egyenlőséget „rontsuk” el egy kicsit, tehát ne várjuk el a tökéletes egyenlőséget a feltételben, csupán annyit, hogy megfelelően közel essen egymáshoz a két érték, azaz

⁵Ez az alapötlet Baek-Brock (1991, 1992) cikkekben már megtalálható, végül ennek egy módosítása „terjedt el” a gyakorlatban

$$\begin{aligned} & \mathbf{P}(|Y_t - Y_s| < \delta, |X_{t-1} - X_{s-1}| < \delta \mid |Y_{t-1} - Y_{s-1}| < \delta) = \\ & = \mathbf{P}(|Y_t - Y_s| < \delta \mid |Y_{t-1} - Y_{s-1}| < \delta) \mathbf{P}(|X_{t-1} - X_{s-1}| < \delta \mid |Y_{t-1} - Y_{s-1}| < \delta). \end{aligned} \quad (24)$$

Ez nyilván némileg torzítja az eredményeket (természetesen $\delta \rightarrow 0$ határértékben visszkapjuk a pontos H_0 -t). A fenti feltételes valószínűségeket visszaírjuk egyszerűbb formába, és ezzel meg is kapjuk a Hiemstra–Jones-teszt nullhipotézisét

$$\begin{aligned} H_0^\dagger : \quad & \mathbf{P}(|Y_t - Y_s| < \delta \mid |Y_{t-1} - Y_{s-1}| < \delta, |X_{t-1} - X_{s-1}| < \delta) = \\ & = \mathbf{P}(|Y_t - Y_s| < \delta \mid |Y_{t-1} - Y_{s-1}| < \delta). \end{aligned} \quad (25)$$

Ezt a kifejezést már lehet a korrelációs integrálok alapuló módszerrel tesztelni, először is átírjuk a feltételes valószínűségeket, így a következőhöz jutunk:

$$\begin{aligned} & \frac{\mathbf{P}(|Y_t - Y_s| < \delta, |Y_{t-1} - Y_{s-1}| < \delta, |X_{t-1} - X_{s-1}| < \delta)}{\mathbf{P}(|Y_{t-1} - Y_{s-1}| < \delta, |X_{t-1} - X_{s-1}| < \delta)} = \\ & = \frac{\mathbf{P}(|Y_t - Y_s| < \delta, |Y_{t-1} - Y_{s-1}| < \delta)}{\mathbf{P}(|Y_{t-1} - Y_{s-1}| < \delta)}. \end{aligned} \quad (26)$$

Felhasználva a korrelációs integrál definícióját, adódik hogy

$$\frac{C_{Y_t, Y_{t-1}, X_{t-1}}(\delta)}{C_{Y_{t-1}, X_{t-1}}(\delta)} = \frac{C_{Y_t, Y_{t-1}}(\delta)}{C_{Y_{t-1}}(\delta)}. \quad (27)$$

Vegyük észre, hogy ugyanerre a kifejezésre jutunk, ha a korrelációs integrált „naiv” módon használjuk, tehát (99) egyenletben egyszerűen kicseréljük a sűrűségfüggvényeket a hozzájuk tartozó korrelációs integrálra.

A következőkben vázlatosan ismertetjük a teszt menetét, idáig viszonylag könnyű dolgunk volt a nullhipotézisek felírásával, természetesen a tesztfüggvények – asszimptotikus – eloszlásának levezetése ennél sokkal bonyolultabb, az eredeti cikkekben megtalálhatóak. Eddig az y_t , y_{t-1} és x_{t-1} tetszőleges valós számokat jelöltek, a következőkben azonban egy véges minta elemeit fogjuk érteni alattuk, ahol $t = 1, \dots, T$. Először is írjuk át (27) egyenletben szereplő elméleti korrelációs integrálokat azok becslőfüggvényeire:

$$\begin{aligned} \hat{C}_{Y_t, Y_{t-1}, X_{t-1}} &= \frac{1}{(2\delta)^3} \frac{1}{T(T-1)} \sum_{t=1}^T \sum_{s \neq t} \mathbf{1}_\delta(y_t, y_s) \cdot \mathbf{1}_\delta(y_{t-1}, y_{s-1}) \cdot \mathbf{1}_\delta(x_{t-1}, x_{s-1}) \\ \hat{C}_{Y_{t-1}, X_{t-1}} &= \frac{1}{(2\delta)^2} \frac{1}{T(T-1)} \sum_{t=1}^T \sum_{s \neq t} \mathbf{1}_\delta(y_{t-1}, y_{s-1}) \cdot \mathbf{1}_\delta(x_{t-1}, x_{s-1}) \\ \hat{C}_{Y_t, Y_{t-1}} &= \frac{1}{(2\delta)^2} \frac{1}{T(T-1)} \sum_{t=1}^T \sum_{s \neq t} \mathbf{1}_\delta(y_t, y_s) \cdot \mathbf{1}_\delta(y_{t-1}, y_{s-1}) \\ \hat{C}_{Y_{t-1}} &= \frac{1}{2\delta} \frac{1}{T(T-1)} \sum_{t=1}^T \sum_{s \neq t} \mathbf{1}_\delta(y_{t-1}, y_{s-1}), \end{aligned} \quad (28)$$

ekkor H_0^\dagger fennállása esetén aszimptotikusan igaz lesz, hogy

$$\lim_{T \rightarrow \infty} \frac{\frac{\hat{C}_{Y_t, Y_{t-1}, X_{t-1}} - \hat{C}_{Y_t, Y_{t-1}}}{\hat{C}_{Y_{t-1}, X_{t-1}} - \hat{C}_{Y_{t-1}}} \sim \mathcal{N}(0, 1), \quad (29)$$

$$\frac{S_{\text{HJ}}(T)}{\sqrt{T}}$$

ahol $S_{\text{HJ}}(T)$ becslésére a Szerzők (Hiemstra–Jones, 1994) a cikk függelékében térnek ki. A szimulációk során ezt az Z -statisztikát számoltuk ki, majd kétoldali p -értéket határoztuk meg hozzá.

3.1.2 Diks–Panchenko-teszt

Diks és Panchenko szerint a Hiemstra és Jones tesztje sokszor jelez okságot indokolatlan esetben is (Diks–Panchenko, 2005), ami elsősorban a iménti szakaszban bemutatott torzító hatás miatt van, ennek minimalizálása motiválja a szerzőket.

Az ötletük, hogy a nullhipotézis fennállása mellett egy súlyfüggvénnyel szorozzák meg (99) egyenlet mindkét oldalát. Világos, hogy tetszőleges $g(\cdot)$ függvénnyel megszorozva a nullhipotézist az nem változik, tehát ezt megtehetjük, de természetesen újra ki kell számolni, hogy immáron milyen eloszlást fog követni a próbafüggvény. A szerzők az $f_{Y_{t-1}}^2(y_{t-1})$ függvényt választották, így a Diks–Panchenko-teszt nullhipotézise

$$H_0^\dagger : \quad \begin{aligned} f_{Y_t, Y_{t-1}, X_{t-1}}(y_t, y_{t-1}, x_{t-1}) f_{Y_{t-1}}(y_{t-1}) &= \\ &= f_{X_{t-1}, Y_{t-1}}(x_{t-1}, y_{t-1}) f_{Y_t, Y_{t-1}}(y_t, y_{t-1}). \end{aligned} \quad (30)$$

Ennek az állításnak is korrelációs integrálon keresztül történik a tesztelése, ahogy azt az előző szakaszban a Hiemstra–Jones-tesztnél láttuk. Felhasználva a (28) jelöléseit az alábbi próbafüggvényhez jutunk:

$$\lim_{T \rightarrow \infty} \frac{\hat{C}_{Y_t, Y_{t-1}, X_{t-1}} \hat{C}_{Y_{t-1}} - \hat{C}_{Y_{t-1}, X_{t-1}} \hat{C}_{Y_t, Y_{t-1}}}{S_{\text{DP}}(T)} \sim \mathcal{N}(0, 1), \quad (31)$$

$$\frac{\sqrt{T}}$$

ahol az $S_{\text{DP}}(T)$ kiszámításáról Diks–Panchenko (2005) cikk A.1 függelékében térnek ki a szerzők. A HJ-teszthez hasonlóan itt is kétoldali p -értéket számoltunk a tesztfüggvény értékéből. Mindkét teszt esetében felmerül a kérdés, hogy milyen δ értékkel számoljunk, ennek optimális értékéről, illetve levezetéséről részletesen (Diks–Panchenko, 2005) cikkben olvashatunk.

Összefoglalóan elmondhatjuk a korrelációs integrálon alapuló módszerekről, hogy a feltételes függetlenség tesztelésének problémáját ötletesen visszavezeti feltétel nélküli eloszlásokra, így folytonos változókra is kezelhetővé válik a probléma. Ugyanakkor a feltételes eloszlásokban a feltétel „gyengítése” azzal jár, hogy nem pontosan az eredeti problémát tudjuk tesztelni segítségünkkel.

3.2 Kopulákkal kapcsolatos tesztek

A kopulák használata mára már teljesen általánossá vált a pénzügyi matematika legtöbb területén (kockázatmenedzsment, árazási modellek stb.), használatukat a valószínűségi változók együttmozgását mérő korreláció hiányosságai tették indokolttá. A pénzügyi piacokon az Eszközök árfolyamai/hozamai szövevényes módon függenek egymástól, ezeket a függőségeket nem lehet lineáris keretek között mérni, ahogyan azt a korreláció teszi. Ugyanakkor az együttes eloszlásukat nem ismerjük, a kopulák ezt a hiányt próbálják kisebb-nagyobb sikerrel pótolni. A bemutatni kívánt módszerhez a kopuláknak csupán néhány tulajdonsága fontos, így viszonylag rövid bevezetőt adunk róluk, magyar nyelven (Varga, 2004) cikkét ajánljuk a részletesebb megismerésükhöz.

A továbbiakban nekünk csupán 2 és 3 dimenziós kopulákra lesz szükségünk, így az egyszerűség kedvéért most 3 dimenziósra mondjuk ki a definíciót valamint mutatjuk be a számunkra szükséges tulajdonságokat, abból egyszerűen lehet szűkíteni 2 változóra.

4. Definíció. *Egy $C: [0, 1]^3 \rightarrow [0, 1]$ eloszlásfüggvény kopula, ha minden változója szerinti marginális eloszlása standard egyenletes.*

Tekinthetjük úgy is, hogy a kopula egy $[0, 1]$ -es értelmezési tartományra átskálázott eloszlásfüggvény, s így tudunk belőle a $[0, 1]$ intervallumba eső véletlen számokat generálni, amiből inverz eloszlásfüggvényekkel véletlen értékeket kapunk. A kopulákkal kapcsolatos legfontosabb tétel következik (szintén három változós formában):

1. Tétel (Sklar, 1959). *Minden 3 változós F eloszlásfüggvényhez létezik olyan, szintén 3 változós C kopula, hogy*

$$C(F_X(x), F_Y(y), F_Z(z)) = F(x, y, z).$$

Tehát, ha a priori a kopulából, vagy a marginálisokból, vagy az együttes eloszlásfüggvényből ismerünk kettőt (vagy feltesszük, hogy ismert), akkor a harmadikat ki tudjuk számolni. A tétel jelentősége abban áll, hogy empirikus problémáknál a marginálisokat tipikusan ismerjük, hiszen az egyes változók empirikus eloszlásfüggvényét a mintából meg tudjuk határozni. Ezek után valamilyen a priori feltevés útján kiválasztjuk a használatos kopulát, s ki tudjuk számolni az együttes eloszlásfüggvényt, illetőleg tudunk generálni az együttes eloszlásfüggvény segítségével mesterséges mintát, amit különböző szimulációkhoz felhasználhatunk.

A kopuláknak számos hasznos és fontos tulajdonsága van, számunkra most a kopulasűrűség, ami kiemelt szerepet játszik. Deriváljuk mindhárom változó szerint a (32) egyenlet mindkét oldalát, így a következőt kapjuk:

$$\frac{\partial^3}{\partial u \partial v \partial w} C(u, v, w) f_X(x) f_Y(y) f_Z(z) = f(x, y, z), \quad (33)$$

ahol $u = F_X(x)$, $v = F_Y(y)$ és $w = F_Z(z)$ valamint f_X, f_Y, f_Z a megfelelő marginálisokhoz és f az együttes eloszlásfüggvényhez tartozó sűrűségfüggvények.

5. Definíció. Egy 3 változós C kopulához tartozó kopulasűrűség alatt a következő függvényt értjük

$$c(u, v, w) = \frac{\partial^3}{\partial u \partial v \partial w} C(u, v, w). \quad (34)$$

Amíg a kopula az eloszlásfüggvényhez kapcsolódó fogalom volt, addig a kopulasűrűség a sűrűségfüggvényhez, tulajdonságai is ehhez hasonlítanak: nem-negatív (de nem biztos hogy egynél kisebb) és a kopulához hasonlóan az egység kockán (négyzeten) van értelmezve. A rövid módszertani áttekintés után két eljárást mutatunk be, melyek segítségével a feltételes függetlenség problémája kezelhető.

3.2.1 Bernstein-kopulán alapuló teszt

Taamoutia et al. (2014) cikke nyomán a kopulán alapuló okság teszt ötlete, hogy a peremeloszlások ismeretében egy kopulából származtatjuk az együttes eloszlást. Azaz az együttes eloszlásokat, pontosabban azok sűrűségfüggvényeit kicseréljük kopulasűrűségekre, s az együttes eloszlás helyett a becslt kopulákat fogjuk összehasonlítani, ebből vezetjük le a tesztet. Ha tehát fennáll (16), akkor igaz, hogy

$$H_0 : f_{Y_t|Y_{t-1}, X_{t-1}}(y_t | y_{t-1}, x_{t-1}) = f_{Y_t|Y_{t-1}}(y_t | y_{t-1}). \quad (35)$$

Először a feltételes sűrűségeket írjuk át, így a H_0

$$\frac{f_{Y_t, Y_{t-1}, X_{t-1}}(y_t, y_{t-1}, x_{t-1})}{f_{Y_{t-1}, X_{t-1}}(y_{t-1}, x_{t-1})} = \frac{f_{Y_t, Y_{t-1}}(y_t, y_{t-1})}{f_{Y_{t-1}}(y_{t-1})}, \quad (36)$$

beírva a kopulasűrűségeket a (33) egyenlet alapján, és bevezetve a $w_t = F_{Y_t}(y_t)$ illetve $u_t = F_{X_t}(x_t)$ jelöléseket

$$\begin{aligned} & \frac{c_{Y_t, Y_{t-1}, X_{t-1}}(w_t, w_{t-1}, u_{t-1}) f_{Y_t}(y_t) f_{Y_{t-1}}(y_{t-1}) f_{X_{t-1}}(x_{t-1})}{c_{Y_{t-1}, X_{t-1}}(w_{t-1}, u_{t-1}) f_{Y_{t-1}}(y_{t-1}) f_{X_{t-1}}(x_{t-1})} = \\ & = \frac{c_{Y_t, Y_{t-1}}(w_t, w_{t-1}) f_{Y_t}(y_t) f_{Y_{t-1}}(y_{t-1})}{f_{Y_{t-1}}(y_{t-1})}, \end{aligned} \quad (37)$$

így egyszerűsítések után az átírt – immáron csak kopulasűrűségeket tartalmazó – nullhipotézis:

$$H_0^* : \frac{c_{Y_t, Y_{t-1}, X_{t-1}}(w_t, w_{t-1}, u_{t-1})}{c_{Y_t, Y_{t-1}}(w_t, w_{t-1}) c_{Y_{t-1}, X_{t-1}}(w_{t-1}, u_{t-1})} = 1. \quad (38)$$

Az világos, hogy a mintából ismerjük az empirikus peremeloszlásokat, már „csak” egy megfelelő empirikus kopulát kell kiválasztani, amivel meg tudjuk becsülni a fenti kifejezést, és meg tudjuk határozni, milyen eloszlást követ.

Idősoros modellekben leggyakrabban alkalmazott kopula a Bernstein-kopula (részletesen lásd Sancetta–Satchell, 2004). A teljesen általános definícióját nem mondjuk ki, már csak a speciálisan erre a feladatra felírandó 2- és 3-dimenziós esetekkel foglalkozunk (általában sem nehezebb megadni a kopulát,

ezért is tekintünk el az általánosságtól). Ehhez tekintsük az (Y_t, Y_{t-1}, X_{t-1}) – ismeretlen – eloszlásból származó $\{(y_t, y_{t-1}, x_{t-1})\}_{t=1}^T$ mintát. Legyenek továbbá $u_t = F_{X_t, T}(x_t)$ és $w_t = F_{Y_t, T}(y_t)$, ahol $F_{Y_t, T}$ és $F_{X_t, T}$ a T hosszú mintából származó empirikus eloszlásfüggvényt jelöli, azaz a mintabeli x_t és y_t értékek az u_t és w_t -hez tartozó empirikus kvantilisek.

Vezessük be továbbá a következő 2- és 3-dimenziós kockákat:

$$B_k^{i_1, i_2} = \left[\frac{i_1 - 1}{k}, \frac{i_1}{k} \right) \times \left[\frac{i_2 - 1}{k}, \frac{i_2}{k} \right), \quad (39)$$

$$B_k^{i_1, i_2, i_3} = \left[\frac{i_1 - 1}{k}, \frac{i_1}{k} \right) \times \left[\frac{i_2 - 1}{k}, \frac{i_2}{k} \right) \times \left[\frac{i_3 - 1}{k}, \frac{i_3}{k} \right), \quad (40)$$

valamint legyen $\mathbf{1}_{B_k^{i_1, i_2}}(\cdot, \cdot)$ és $\mathbf{1}_{B_k^{i_1, i_2, i_3}}(\cdot, \cdot, \cdot)$ a kockákba esésre vonatkozó indikátor függvény.

Ekkor a 2-dimenziós, becült Bernstein-kopulákat egy tetszőleges pontban a mintán az alábbi módon számítjuk

$$\begin{aligned} & \hat{c}_{Y_t, T, Y_{t-1}, T}(g_1, g_2) = \\ & = \frac{1}{T} \sum_{t=1}^T \left[k^2 \sum_{i_1=1}^k \sum_{i_2=1}^k \mathbf{1}_{B_k^{i_1, i_2}}(w_t, w_{t-1}) \prod_{\ell=1}^2 \binom{k}{i_\ell - 1} g_\ell^{i_\ell - 1} (1 - g_\ell)^{(k - i_\ell - 1)} \right], \end{aligned} \quad (41)$$

$$\begin{aligned} & \hat{c}_{Y_{t-1}, T, X_{t-1}, T}(g_2, g_3) = \\ & = \frac{1}{T} \sum_{t=1}^T \left[k^2 \sum_{i_1=1}^k \sum_{i_2=1}^k \mathbf{1}_{B_k^{i_1, i_2}}(w_{t-1}, u_{t-1}) \prod_{\ell=1}^2 \binom{k}{i_\ell - 1} g_\ell^{i_\ell - 1} (1 - g_\ell)^{(k - i_\ell - 1)} \right], \end{aligned} \quad (42)$$

továbbá a 3-dimenziókat az alábbi módon

$$\begin{aligned} & \hat{c}_{Y_t, T, Y_{t-1}, T, X_{t-1}, T}(g_1, g_2, g_3) = \\ & = \frac{1}{T} \sum_{t=1}^T \left[k^3 \sum_{i_1=1}^k \sum_{i_2=1}^k \sum_{i_3=1}^k \mathbf{1}_{B_k^{i_1, i_2, i_3}}(w_t, w_{t-1}, u_{t-1}) \prod_{\ell=1}^3 \binom{k}{i_\ell} g_\ell^{i_\ell} (1 - g_\ell)^{(k - i_\ell - 1)} \right]. \end{aligned} \quad (43)$$

Ebből a három – a mintán megbecsült – kopulából egy ún. *oksági mérték*et számolunk ki⁶

$$\text{CM}_T = \frac{1}{T} \sum_{t=1}^T \log \left(\frac{\hat{c}_{Y_t, T, Y_{t-1}, T, X_{t-1}, T}(w_t, w_{t-1}, u_{t-1})}{\hat{c}_{Y_t, T, Y_{t-1}, T}(w_t, w_{t-1}) \cdot \hat{c}_{Y_{t-1}, T, X_{t-1}, T}(w_{t-1}, u_{t-1})} \right). \quad (44)$$

Ez mutató Taamoutia et al. (2014) alapján már aszimptotikusan normális eloszlást fog követni a megfelelő várható értékkel és szórással:

$$\lim_{T \rightarrow \infty} T k^{-3/2} \frac{2\text{CM}_T - T^{-1} k^{3/2} \xi}{\sigma} \sim \mathcal{N}(0, 1),$$

⁶Geweke (1982) cikkben szereplő oksági mérték nem-lineáris kiterjesztése.

ahol $\sigma = \sqrt{2}(\pi/4)^{3/2}$ valamint $\xi = -\pi^{3/2}/8 + \pi/2k^{-1/2} - k^{-1}(\pi^{1/2} - 1)$.

Ahogy a korrelációs integrál alapú teszteknel, itt is el kell látni értékkel egy „bandwidth” paramétert, a k -t, ami jelen esetben a kockák méretében jelenik meg. A szerzők nyomán ezt $k = \lfloor T^{1/2} \rfloor$ nagyságúra választjuk.

3.2.2 Független-kopulán alapuló teszt

Az utolsó módszer, ami bemutatásra kerül, egy általunk kidolgozott teszt a nem-paraméteres okság vizsgálatára. Módszerünk alapja, hogy két folytonos skálán mért változó – gondolhatunk itt rögtön idősorokra – feltétel nélküli függetlenségének tesztelése kopulák segítségével megoldott feladat (lásd. Genest–Remillard, 2004; Genest et al., 2007). Ugyanakkor számunkra a feltételes függetlenség, ami igazán érdekes, hiszen éppen a feltétel ismeretében tudjuk kiszűrni azokat a hatásokat, melyek nem közvetlenül lépnek fel Y_t és X_{t-1} között.

A probléma hasonló mint a korábbi korrelációs integrálon alapuló módszer bemutatásnál, a folytonos skálán mért idősori értékeknél véges mintán lehetetlen egy-egy adott pontra elkészíteni a feltételes eloszlást, mert a rész-minta egy elemből fog állni. Ennek áthidalására – Hiemstra és Jones ötletétől vezérelve – beosztottuk a mintában szereplő y_{t-1} értékeket k darab csoportba, így kaptunk k számú egymástól elszeparált almintát. Ezt követően minden almintán elvégeztük a kopula alapú függetlenség tesztet, amit \mathbf{R} környezetben a *copula* csomag tartalmaz (Hofert et al. 2015), ami szintén Genest–Remillard módszerét használja. Majd az így kapott p -értékeket aggregáltuk, a módszer Fisher (1948) ötletén alapul, aki azonban független rész-mintákra dolgozta ki a módszert. Nem független minták esetére Brown (1975) adott egy aggregálási módszert. Ez azonban viszonylag nehézkesen alkalmazható a gyakorlatban, ugyanis numerikus integrálást is tartalmaz, s így a számítási idő rendkívüli módon megnő, a szimuláció számítási ideje egyszerűen túl hosszú. Alternatív megoldást kínál Poole et al. (2015), az ő módszerüket – amire empirikus Brown-módszer néven fogunk hivatkozni ezentúl – azonban némileg módosítani kellett (melyről később lesz szó), hogy esetünkben is alkalmazni lehessen.

Összefoglalva – és formalizálva – a fentieket, a következő lépéseket kell követni ez eljárás során:

1. Definiáljuk a feltételi halmazokat: legyen $k > 1$ és

$$B_k^i = \left[\frac{i-1}{k}, \frac{i}{k} \right),$$

ahol $i \in \{1, \dots, k\}$, így természetesen $\cup_{i=1}^k B_k^i = [0, 1)$ és $B_k^i \cap B_k^j = \emptyset$.

2. Bontsuk szét a teljes mintánkat az alábbi almintákra

$$(y_t^i, x_{t-1}^i) := \{(y_t, x_{t-1}) \mid F_{Y_{t-1}}(y_{t-1}) \in B_k^i\},$$

és az empirikus marginálisokra vezessük be az $F_{Y_t, T}^i(\cdot)$ és $F_{X_{t-1}, T}^i(\cdot)$ jelöléseket. Legyenek továbbá az egyes alminták idő indexei a \mathcal{T}_k^i -vel

jelölve, ezekre természetesen igaz, hogy $\cup_{i=1}^k \mathcal{T}_k^i = \{1, \dots, T\}$ és $\mathcal{T}_k^i \cap \mathcal{T}_k^j = \emptyset$. Ez utóbbi halmazok hosszát pedig jelöljük rendre T_k^i -vel, ekkor természetesen $\sum_{i=1}^k T_k^i = T$.

3. Minden S_k^i részmintán végezzük el a függetlenség tesztet. Ehhez először becsüljük meg az empirikus kopulákat a részmintákon:

$$\hat{C}_T^i(u, w) = \frac{1}{T_k^i} \sum_{t \in \mathcal{T}_k^i} \mathbf{1}(F_{Y_{t,T}}^i(y_t) \leq u) \mathbf{1}(F_{X_{t-1,T}}^i(x_{t-1}) \leq w) .$$

A nullhipotézisben a függetlenséget a független kopula segítségével tudjuk megfogalmazni:

$$H_0^i : \hat{C}_T^i(u, w) = u \cdot w .$$

Magában a függetlenség tesztben a próbafüggvény az empirikus kopula pontonkénti eltérését vizsgálja a független kopulától:

$$\text{CM}(u, w) = \sqrt{n}(\hat{C}_T^i(u, w) - u \cdot w) .$$

Ez a kifejezés véletlen u, w értékekre aszimptotikusan Cramer–von Mises statisztikát követ⁷, melynek saját táblázatából olvassuk ki az i -edik almintához tartozó teszt p -értékét, jelöljük ezt p^i -vel. Megjegyezzük, hogy ezt a lépést teljes egészében a **copula** csomag végzi.

4. Aggregáljuk a p -értékeket az empirikus Brown-módszer segítségével:

$$\Psi = -2 \sum_{i=1}^k \log p^i ,$$

ami egy átskálázott χ^2 eloszlást követ, azaz $\Psi \sim c\chi_{2f}^2$. A c és f konstansok kiszámítását az F-2 Függelékben ismertetjük. Ezek alapján a „globális” p -értéket a következőképpen számolhatjuk ki

$$p = 1 - F_{\chi_{2f}^2}(\Psi/c) ,$$

amennyiben ez az előre rögzített szignifikanciaszint alatt van, akkor elvetjük az eredeti, feltételes függetlenségre vonatkozó nullhipotézist.

A módszer fontos része a B_k^i -k halmazok rendszerének meghatározása, természetesen nem kötelező egyenletesen k részre osztani a $[0, 1]$ intervallumot, viszont minden esetben biztosítani kell az elegendő mintaelemszámot a részmintákban a számításokhoz. További vizsgálatra lehet érdemes a B_k^i -k valamilyen más – bizonyos szempontból optimális – felosztása, ez nem tárgya ennek a cikknek.

⁷A Cramer–von Mises statisztika eloszlások közti távolságot mér, ebből a szempontból hasonló a Kolmogorov–Szmirnov statisztikához, részletesebben Csörgő–Faraway (1996) tanulmányát ajánljuk.

A kopula alapú módszerekről összefoglalásképpen elmondhatjuk, hogy a tesztelendő feltételes függetlenséget kopulák függetlenségére vezetik vissza. Ennek a megközelítésnek az előnye, hogy nem a teljes \mathbb{R}^2 és \mathbb{R}^3 halmazokon kell vizsgálni sűrűségfüggvényeket, hanem az egység oldalú négyzeteken, kockákon ezzel gyakorlatilag besűrítjük a mintapontokat. Általában a kopulák illesztése szoftveresen viszonylag jól támogatott, ugyanakkor ezek konkrét problémákra való alkalmazása, ahogy az esetünkben is történik, már önálló programozási munkát igényel.

Ezzel a végére értünk a tesztek – legalábbis vázlatosan – bemutatató résznek, amiből kiderült, hogy a nem-paraméteres okság tesztelése alapvetően egy feltételes függetlenség tesztre vezethető vissza. Ezt pedig eltérő metodológiákkal próbálják kezelni a szerzők, azonban közös vonása a teszteknek, hogy relatíve nagy a számításigényük, és mivel asszimptotikus tulajdonságaikat ismerjük a próbafüggvényeknek, így elengedhetetlen a megfelelő mintanagyság az alkalmazásukhoz.

4 Szimulációs példa

A számításokhoz R környezetet használtunk, elsősorban a meggyőző teljesítménye miatt, illetve a kopulák illesztése csomagokkal jól támogatott. A korábban bemutatott négyféle nem-paraméteres Granger-tesztből a Hiemstra–Jones- és Diks–Panchenko-teszteket egy külső szoftver segítségével futtattuk.⁸ A Bernstein-kopulán alapuló tesztet teljes egészében R nyelven implementáltuk, a független kopula módszer lényegi részét a `copula` csomag végzi (Hofert et al., 2015) illetve (Poole et al., 2015) módszert a szerzők által írt függvények, valamint a referencia tesztként szolgáló klasszikus Granger-tesztet többek között az `lmtest` csomag (Zeileis–Hothorn, 2002) tartalmazza.

A szimulációkhoz négy különböző adatgeneráló folyamatot használtunk, melyek az alábbiak:

$$\text{DGP1:} \quad Y_t = \alpha Y_{t-1} + \beta X_{t-1} + \epsilon_t^Y, \quad X_t = \gamma X_{t-1} + \epsilon_t^X, \quad (46)$$

$$\text{DGP2:} \quad Y_t = \alpha Y_{t-1} + \beta X_{t-1}^2 + \epsilon_t^Y, \quad X_t = \gamma X_{t-1} + \epsilon_t^X, \quad (47)$$

$$\text{DGP3:} \quad Y_t = \beta Y_{t-1} X_{t-1} + \epsilon_t^Y, \quad X_t = \gamma X_{t-1} + \epsilon_t^X, \quad (48)$$

$$\text{DGP4:} \quad Y_t = \sigma_t \epsilon_t^Y, \quad \sigma_t^2 = \alpha + \beta X_{t-1}^2, \quad X_t = \gamma X_{t-1} + \epsilon_t^X, \quad (49)$$

ahol $\epsilon_t^X, \epsilon_t^Y \sim \mathcal{N}(0, 1)$ független valószínűségi változók. Valamennyi adatgeneráló folyamatot úgy definiáltuk, hogy $\beta \neq 0$ esetben $X \rightarrow Y$ okság álljon fenn, valamint $\alpha, \beta, \gamma < 1$ feltétel mellett mind a négy folyamat stacioner, így a próbák megbízható eredményt szolgáltatnak.⁹

⁸Ez gyakorlatilag egy, a szerzőpáros által írt C kód, amit apróbb módosítások után újrafordítottunk és R-ből hívtuk meg a programot, majd visszaolvastuk az eredményeket R-be, a kód megtalálható a Cees Diks weblapján: <http://www.uva.nl/en/about-the-uva/organisation/staff-members/content/d/i/c.g.h.diks/c.g.h.diks.html>

⁹Taamoutia et al. (2014) is többek között ezeket a DGP-eket használják a szimulációkban.

Valamennyi adatgeneráló folyamatnál „ok” szerepét betöltő X_t idősor elsőrendű autoregresszív idősor, a különbség az „okozat” idősor formájában van. A DGP1 egyszerű lineáris modell gyakorlatilag referenciaként szolgál, a DGP2-ben az ok kvadratikus formában jelenik meg, a DGP3 folyamat felfogható egy változó együtthatós elsőrendű autoregresszív modellként, ahol az együttható szintén egy AR(1)-ből származik, míg a DGP4 leginkább egy GARCH modellre hasonlít, amiben a feltételes variancia egy exogén változótól függ. Az elsőt leszámítva mindegyikben van valamilyen nemlinearitás, így azt várjuk, hogy esetükben a Granger-féle teszt „rossz” eredményt szolgáltat, ugyanakkor a másik három módszer jelezzen a $\beta \neq 0$ esetben.

A szimulációk során 250 periódus hosszú idősorokat generáltunk ezt praktikusán úgy értelmezzük, mintha egy éves időhorizonton, napi frekvencián vizsgálnánk tőzsdei idősorokat (pl. napi záróárfolyamot). Ezekre a generált idősorokra minden esetben alkalmazzuk a teszteket és a kapott p -érték valamint az előre rögzített szignifikancia szint alapján döntünk a nullhipotézisről. Majd a szimuláció végén minden különböző β értékre ($\beta = 0$ esetben ez a H_0 , egyébként a H_1 feltétel teljesülését jelenti) kiszámoltuk milyen arányban vetettük el a nullhipotézist, azaz néhány pontban kiszámoltuk a tesztek erőfüggvényeit, kiegészítve azzal az esettel, amikor a nullhipotézis igaz. Ez utóbbi feltétel mellett ideális esetben vissza kellene kapni az előre rögzített szignifikanciaszintet, minden egyéb β érték mellett pedig elméletileg minden esetben el kellene vetni a nullhipotézist. Az eredményeket bemutatásakor „HJ”-vel a Hiemstra–Jones-tesztet, „PD”-vel a Panchenko–Diks-tesztet, „CB”-vel a Bernstein-kopulán, „CI”-vel a független kopulán alapuló tesztet, míg „GR”-rel a Granger-féle tesztet jelöltük.

Az 1. táblázat a 0 és 0.5 közötti β értékek esetében vizsgálja a tesztek teljesítményét $\alpha = \gamma = 0.5$. Ahogyan vártuk a DGP1 esetében (lineáris modell), a Granger-féle teszt érzékenyebben reagál a β növekedtével mint a többi, nem-paraméteres társa. Nem sokkal elmaradva tőle a független kopula alapú teszt is viszonylag alacsony β mellett nagy arányban jelzi az okságot, a két korrelációs integrálon alapuló módszer csupán 0.3-as β érték körül mutat szignifikáns okságot, a legrosszabb eredményt pedig a Bernstein-kopulán alapuló teszt produkálja. Ugyanakkor a többi – nem lineáris – eset mindegyikében nagyon rosszul teljesít a Granger-féle teszt, hiszen rendkívül kicsi arányban jelez okságot, itt már csak a nem-paraméteres próbákban bízhatunk. A DGP2 (kvadratikus) esetében a korrelációs integrálon alapuló tesztek bizonyultak hatékonyabbnak, míg a DGP3-nál (változó paraméterű AR) a kopula alapúak reagáltak gyorsabban a β paraméter növekedtére. A DGP4 esetében (GARCH jellegű) a korrelációs integrál alapú tesztek jeleztek korábban, ráadásul a független kopula alapú teszt az adott paraméter tartományban nagyon rossz arányban jelzett okságot.

A szimulációk tanulsága, hogy a nem-paraméteres próbák esetében sem mindegy, hogy milyen jellegű nem-linearitás jelenik meg a folyamatban. A kopula alapú módszerek érzékenyebbek a várható értékben történő elmozdulásokra (DGP2, DGP3), míg a korrelációs integrálon alapuló módszerek jobban teljesítenek a második momentumban bekövetkező hatásokra.

	β	HJ	PD	CB	CI	GR
DGP1	0.00	0.041	0.029	0.001	0.040	0.050
	0.10	0.080	0.073	0.009	0.114	0.436
	0.20	0.308	0.373	0.052	0.461	0.952
	0.30	0.743	0.816	0.261	0.837	0.999
	0.40	0.968	0.984	0.708	0.983	1.000
	0.50	0.999	0.999	0.931	1.000	1.000
DGP2	0.00	0.045	0.039	0.001	0.051	0.052
	0.10	0.156	0.157	0.014	0.085	0.082
	0.20	0.661	0.680	0.232	0.183	0.163
	0.30	0.948	0.958	0.689	0.370	0.256
	0.40	0.997	0.998	0.970	0.629	0.300
	0.50	1.000	1.000	0.997	0.795	0.369
DGP3	0.00	0.028	0.020	0.055	0.039	0.053
	0.10	0.045	0.038	0.087	0.086	0.061
	0.20	0.139	0.125	0.309	0.410	0.076
	0.30	0.359	0.361	0.692	0.824	0.113
	0.40	0.730	0.734	0.964	0.986	0.149
	0.50	0.910	0.917	0.996	0.997	0.244
DGP4	0.00	0.037	0.031	0.069	0.051	0.044
	0.10	0.628	0.548	0.191	0.055	0.105
	0.20	0.931	0.892	0.432	0.081	0.158
	0.30	0.991	0.974	0.607	0.093	0.144
	0.40	0.996	0.993	0.780	0.110	0.176
	0.50	0.998	0.998	0.877	0.106	0.187

1. táblázat. $T = 250$, $N = 1000$, $\beta \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$,
nominális szignifikancia szint= 5%

Ez alapján nem gondoljuk, hogy létezik – legalábbis jelenleg – egy olyan módszer, amit önmagában lehetne alkalmazni, a gyakorlati problémáknál célszerű mindegyik módszerrel megvizsgálni az adott idősorokat. A továbbiakban mi is ezt a gyakorlatot fogjuk követni.

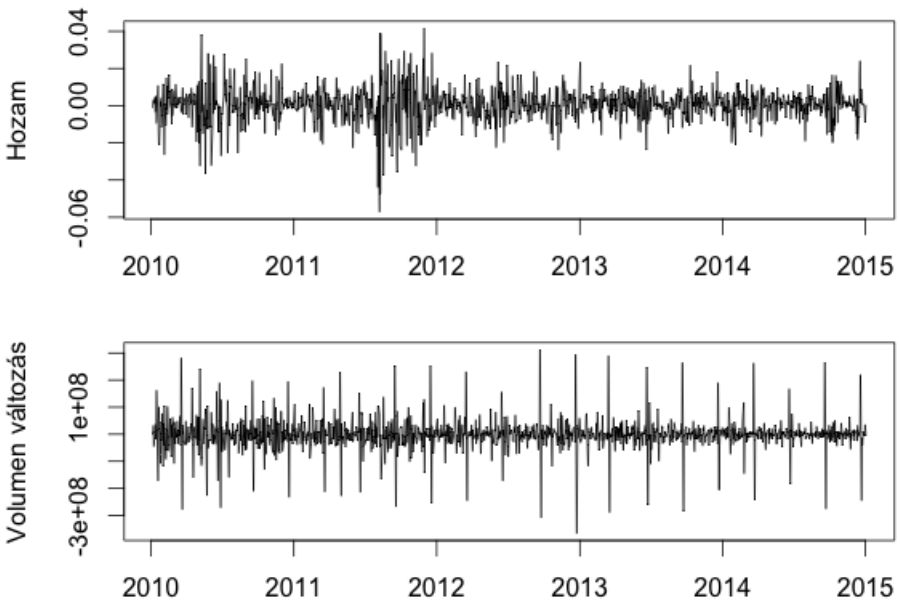
5 Empirikus példa: hozam és kereskedési volumen kapcsolata

A tőzsdei részvények hozama és kereskedési volumenük közötti viszony régóta témája a pénzügyi ökonometriának. Az első cikk, ami ezen a területen készült Granger–Morgenstern (1963) szerzők nevéhez fűződik, ők még csupán heti kereskedési adatok alapján próbálta egyes részvények illetve részvényindexek hozama és volumene közti kapcsolatot feltárni. Azóta természetesen sok más elemzés is napvilágot látott¹⁰, kiváló összefoglalót ad a téma empirikus eredményeiről és az ezeket magyarázni próbáló elméleti modellekről Karpoff (1987). Általában az mondható el, hogy a hozam volumenre gyakorolt hatását könnyebb kimutatni, hiszen a piaci szereplők nagy része árjelzés alapján kereskedik, a fordított irány ilyen szempontból érdekesebb, hiszen ebben az esetben az sem egyértelmű, hogy létezik-e ilyen hatás.

¹⁰Többek között Hiemstra–Jones (1994) cikk is ezen a példán teszteli az általuk kifejlesztett tesztet

A korábban tárgyalt módszertani részeket bemutató cikkek is szinte kivétel nélkül a hozam-volumen kapcsolatán keresztül illusztrálják elméleti eredményeiket. Hiemstra–Jones (1994) hosszú adatsorokon (30 éves idősor a II. világháború előtt és 40 éves utána) mindkét irányú nem-lineáris kapcsolatot kimutatták a Dow Jones Industrial Average (továbbiakban DJIA) részvény index esetében. Ugyanezen az indexen végzett elemzést Diks–Panchenko (2005), 1950 és 1990 között szintén bizonyították mindkét irányú nemlineáris kauzalitást. Míg Bouezmarni et al. (2009) 1988 és 2005 között vizsgálták az S&P 500-at, ahol mindkét irányban talált nem lineáris kapcsolatokat. Közös továbbá az említette cikkekben, hogy klasszikus – lineáris – okságot csak úgy sikerült kimutatni, ha az oknak a hozamot, okozatnak pedig a volument tekintették, fordított irányban nem.

Elemzésünk során mi a DJIA indexet vizsgáltuk napi frekvencián. Hozam alatt a napi záró indexérték loghozamát értjük, a volumen pedig a részvényindexbe tartozó papírok napi összes kereskedési volumenének első differenciája (darabszámban megadva). Az vizsgálat időtartomány 2010.01.04-től 2015.01.02-ig terjed, tehát 5 év kereskedési napjainak adatai, ezeket láthatjuk a 2. ábrán.

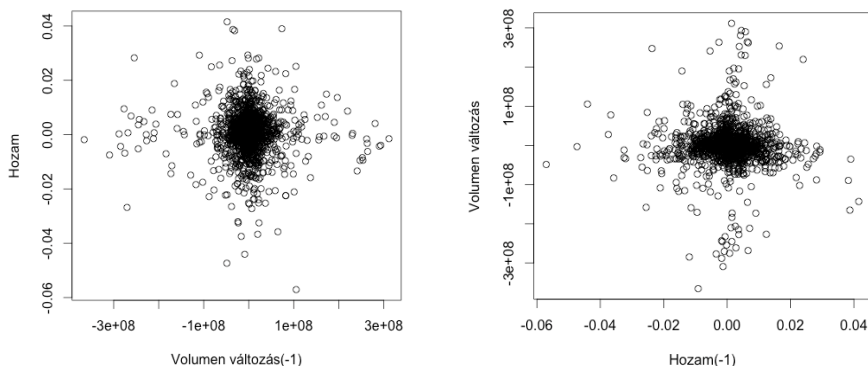


2. ábra. DJIA index hozama és volumen változása 2010.01.04-től 2015.01.02-ig

Mindkét idősor esetben vizsgáltuk a stacionaritást, a loghozam és a volumen első differenciája is egyértelműen stacioner mind a Dickey-Fuller, mind a KPSS teszt alapján. Differencia képzés nélkül a volumen csak Dickey-Fuller teszt alapján tekinthető stacionernek, ezért döntöttünk az első rendű integrálás mellett. Meg kell még jegyezni, hogy a volumen változás esetében negyedévenként tapasztalható kiugrások egyrészt az osztalék fizetésnek, másrészt a szokásos negyedéves jelentésnek köszönhető. Ettől a hatástól nem sérült a

stacionaritás, így amellet döntöttünk, hogy nem szűrjük ki az adatokból, hiszen ez is egyfajta információvesztést eredményezne.

Az oksági viszonyt egy napos késleltetésben vizsgáltuk, ezt pontdiagramon is megpróbáltuk vizualizálni (természetesen ez egy feltétel nélküli eloszlás, így csalóka lehet az ábra), amit a 3. ábrán mutatunk be.



3. ábra. Hozam és volumen változás viszonya a DJIA index esetében 2010.01.04-től 2015.01.02-ig

Az ábrákról egyértelmű tendenciát nem tudtunk leolvasni, „ránézésre” kiugró értékek is gyakorlatilag minden irányban tapasztalhatóak, így semmiféle prekoncepciónk nem volt a tesztek eredményével kapcsolatban. A 2. táblázatban közöljük tesztek p -értékeit.

	$R \Rightarrow \Delta V$	$\Delta V \Rightarrow R$
HJ	0.015**	0.012**
PD	0.016**	0.012**
CB	0.018**	0.014**
CI	0.036**	0.244
GR	0.006***	0.169

2. táblázat. Oksági tesztek eredményei a DJIA indexen, R -rel a hozamot, ΔV -vel a volumen változást jelöltük

Az eredményekből kiderül, hogy 5 éves időhorizonton egyértelműen teljesül, hogy az egy nappal késleltetett hozam oka a másnapi volumenváltozásnak, 5%-on minden teszt szignifikáns hatást mutat, a Granger-féle okságteszt pedig 1%-on is szignifikáns hatást jelez. Fordítva ugyanakkor a klasszikus Granger-teszt valamint a független kopulán alapuló teszt nem mutat okságot, ellentétben a többi teszttel, mely összecseng az irodalom eredményeivel. Utóbbi alapján úgy gondoljuk, a fordított okság nem a várható értékben jelenik meg, ugyanis arra a független kopula alapú teszt érzékeny, hasonlóan

a Granger-féle teszthez (feltéve persze, ha lineáris módon hat a volumen a hozam várható értékére).

6 Összegzés

Munkánkban bemutatunk három ismert, a nem-paraméteres (vagy másképp erős) Granger-okság tesztelésére alkalmas módszert, illetve egy önállóan kidolgozott eljárást. Majd a módszereket összehasonlítottuk szimulációk segítségével különböző adatgeneráló folyamatok esetén, végül egy empirikus példán szemléltettük gyakorlati használhatóságukat.

A szimulációink eredménye alátámasztja a bevezetőben is taglalt jelenséget, miszerint ezek a tesztek méltatlanul alulreprezentáltak ökonometriai vizsgálatokban, ennek oka vélhetőleg a szoftveres támogatás hiánya. Az eredmények láttán megállapítható, hogy a nem-paraméteres tesztek nem teljesítenek sokkal rosszabbul lineáris esetben, mint a klasszikus Granger-teszt, ugyanakkor messze felülmúlják azt nem-lineáris esetben, így mindenképpen érdemes kipróbálni őket, olyan esetben, amikor a Granger-teszt nem jelez okságot, ám az intuíció mégis az ellenkezőjét súgja. A szimulációk során továbbá azt tapasztaltuk, hogy bizonyos esetekben az általunk kidolgozott módszer hatékonyabb az irodalomban eddig ismereteseknél, így indokoltnak találjuk használatát.

Az empirikus példánk a tőzsdei hozamok és a kereskedési volumen közti okságot hivatott feltérképezni, azt találtuk, hogy hosszú távon mindkét irányú okság igazolható, azonban rövidebb távon egyiket sem lehet szignifikánsan kimutatni. Hosszú távon ez az irodalommal összecsengő eredmény, ugyanakkor meglepő a rövid távon kapott eredmény. Ehhez azonban hozzá kell tenni, hogy a nem-paraméteres tesztek kismintás tulajdonságai nem meggyőzőek, elsősorban azért, mert a tesztek valamilyen határeloszlási tételen alapulnak.

További kutatási irányként elsősorban a független kopulán alapuló teszt precíz kidolgozását látjuk, illetve egyéb olyan empirikus példák vizsgálatát, amikben az okság vélhetően valamilyen nem-lineáris módon hat.

Függelék

F1 Az 1. Lemma bizonyítása

1. Lemma. *Amennyiben a (22) igaz, úgy különböző t és s időpontokra*

$$\begin{aligned} & \mathbf{P}(|Y_t - Y_s| < \delta, |X_{t-1} - X_{s-1}| < \delta \mid Y_{t-1} = Y_{s-1} = y) = \\ & = \mathbf{P}(|Y_t - Y_s| < \delta \mid Y_{t-1} = Y_{s-1} = y) \mathbf{P}(|X_{t-1} - X_{s-1}| < \delta \mid Y_{t-1} = Y_{s-1} = y). \end{aligned} \tag{50}$$

Bizonyítás. Ennek belátására írjuk fel a valószínűségeket indikátorfüggvények várható értékeként¹¹, így:

$$\begin{aligned} & \mathbf{P}(|Y_t - Y_s| < \delta, |X_{t-1} - X_{s-1}| < \delta \mid Y_{t-1} = Y_{s-1} = y) = \\ & = \int_{\mathbb{R}^4} \mathbf{1}(|y_t - y_s| < \delta, |x_{t-1} - x_{s-1}| < \delta) f(y_t, y_s, x_{t-1}, x_{s-1} \mid y) d(y_t, y_s, x_{t-1}, x_{s-1}). \end{aligned} \quad (51)$$

Kihasználva, hogy t és s időpontok egymástól független megfigyeléseket indukálnak, szorzatra bontható az együttes sűrűségfüggvény

$$= \int_{\mathbb{R}^4} \mathbf{1}_\delta(y_t, y_s) \mathbf{1}_\delta(x_{t-1}, x_{s-1}) f(y_t, x_{t-1} \mid y) f(y_s, x_{s-1} \mid y) d(y_t, y_s, x_{t-1}, x_{s-1}). \quad (52)$$

Most már alkalmazhatjuk a (22) egyenletben megfogalmazott nullhipotézist, így

$$\begin{aligned} & = \int_{\mathbb{R}^4} \mathbf{1}_\delta(y_t, y_s) \mathbf{1}_\delta(x_{t-1}, x_{s-1}) f(y_t \mid y) f(x_{t-1} \mid y) f(y_s \mid y) f(x_{s-1} \mid y) d(y_t, y_s, x_{t-1}, x_{s-1}) = \\ & = \int_{\mathbb{R}^4} \mathbf{1}_\delta(y_t, y_s) \mathbf{1}_\delta(x_{t-1}, x_{s-1}) f(y_t, y_s \mid y) f(x_{t-1}, x_{s-1} \mid y) d(y_t, y_s, x_{t-1}, x_{s-1}) = \\ & = \int_{\mathbb{R}^2} \mathbf{1}_\delta(y_t, y_s) f(y_t, y_s \mid y) d(y_t, y_s) \int_{\mathbb{R}^2} \mathbf{1}_\delta(x_{t-1}, x_{s-1}) f(x_{t-1}, x_{s-1} \mid y) d(x_{t-1}, x_{s-1}) = \\ & = \mathbf{P}(|Y_t - Y_s| < \delta \mid Y_{t-1} = Y_{s-1} = y) \mathbf{P}(|X_{t-1} - X_{s-1}| < \delta \mid Y_{t-1} = Y_{s-1} = y). \end{aligned} \quad (53)$$

□

F2 Az empirikus Brown-módszer módosítása

A szükséges paraméterek tehát:

$$f = \frac{\mathbf{E}^2 \Psi}{\mathbf{D}^2 \Psi}, \quad c = \frac{k}{f},$$

ahol $\mathbf{E} \Psi = 2k$ és $\mathbf{D}^2 \Psi = 4k + 2 \sum_{i < j} \text{Cov}(-2 \log p^i, -2 \log p^j)$. Ez utóbbi kovarianciának a kiszámítása, pontosabban megbecslése azonban nem triviális feladat, hiszen a p -értékekből nem rendelkezünk mintával. Poole et al. (2015) cikkben éppen ezzel foglalkozik, a szerzőknek sikerült a p -értékek közti kovarianciát a minták közti kovarianciára visszavezetni. Az általuk ehhez készített kódot használjuk fel a szimulációkhoz a későbbiekben.

Esetünkben a p -értékek kétmintás próbából származnak, így kicsit módosítani kell a számításokat. Az alábbi kifejezést igyekszünk megbecsülni

$$\mathbf{D}^2 \Psi = 4k + 2 \sum_{i < j} \text{Cov}(-2 \log p^i, -2 \log p^j). \quad (54)$$

Vezessük be a következő jelölést (a szerzők nyomán), $w_i = 2 \log p_i$, ekkor az elméleti kovarianciát becsülhetjük a szokásos módon

$$\text{Cov}(-2 \log p^i, -2 \log p^j) = \mathbf{E}((w^i - \mathbf{E}w^i)(w^j - \mathbf{E}w^j)).$$

¹¹ $\mathbf{P}(X \in A) = \mathbf{E}(\mathbf{1}(X \in A))$

A probléma, hogy w_i -ből kellene úgy mintát venni, hogy csak egy darab áll rendelkezésre belőle. Ezt oldják fel a szerzők úgy, hogy egyes elemeihez rendelt percentilisek transzformáltjai lesznek a w_i -ből származó mintaelemek. Azaz t -ik mintaelem

$$w_t^i = -2 \log F_{Y_t^i, X_{t-1}^i}(y_t^i, x_{t-1}^i),$$

ahol $F_{Y_t^i, X_{t-1}^i}$ az i -ik almintához tartozó – és a szerzőkkel ellentétben nem egy-, hanem – kétváltozós empirikus eloszlásfüggvény. Az ebből kiszámolt i -edik és j -edik alminták közötti kovarianciát c_{ij} -vel jelölve már ki tudjuk számolni a paramétereket:

$$f = \frac{2k}{4k + 2 \sum_{i < j} c_{ij}}, \quad c = \frac{k}{f}.$$

F3 További szimulációs eredmények

Elsőként a korábbi $T = 250$ hosszú mintát $T = 100$ -ra rövidítettük, ennek a szimulációnak az eredményeit láthatjuk a 3. táblázatban.

	β	HJ	PD	CB	CI	GR
DGP1	0.00	0.034	0.025	0.000	0.038	0.064
	0.10	0.048	0.037	0.000	0.094	0.204
	0.20	0.139	0.140	0.000	0.212	0.593
	0.30	0.324	0.365	0.007	0.461	0.892
	0.40	0.617	0.668	0.040	0.751	0.989
	0.50	0.863	0.890	0.151	0.913	0.999
DGP2	0.00	0.031	0.021	0.000	0.044	0.053
	0.10	0.073	0.067	0.001	0.065	0.077
	0.20	0.252	0.253	0.002	0.121	0.153
	0.30	0.573	0.582	0.028	0.211	0.219
	0.40	0.798	0.804	0.105	0.303	0.321
	0.50	0.931	0.937	0.274	0.428	0.348
DGP3	0.00	0.024	0.013	0.002	0.048	0.068
	0.10	0.029	0.024	0.000	0.081	0.064
	0.20	0.062	0.042	0.008	0.179	0.051
	0.30	0.150	0.130	0.033	0.447	0.098
	0.40	0.323	0.308	0.148	0.696	0.159
	0.50	0.540	0.537	0.397	0.926	0.205
DGP4	0.00	0.025	0.018	0.001	0.044	0.053
	0.10	0.267	0.202	0.001	0.057	0.098
	0.20	0.498	0.404	0.012	0.076	0.123
	0.30	0.677	0.565	0.019	0.067	0.147
	0.40	0.785	0.706	0.045	0.087	0.159
	0.50	0.872	0.792	0.065	0.091	0.186

3. táblázat. $T = 100$, $N = 1000$, $\beta \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$,
nominális szignifikancia szint = 5%

Az eredményekből jól látható, hogy a tesztek jelentősen veszítenek erejükből, ahogy a minta egyre rövidül. Ez elsősorban annak köszönhető, hogy minden nem-paraméteres teszt valamilyen határeloszlási tételre épít. Továbbá

az is látszik, hogy az „erőviszonyok” érdemben nem változnak: a független kopula alapú teszt továbbra is a DGP1 és DGP3 esetben teljesít jobban. Ugyanakkor a Bernstein-kopulán alapuló módszer ereje jelentősen visszaesik a minta rövidülésével.

A második konfigurációban $T = 500$ hosszú idősorokat generáltunk és ezen vizsgáltuk a teszteket, az eredményeket a 4. táblázat tartalmazza.

Az eredményekből látható, hogy elsősorban a Bernstein-kopulán alapuló módszer, ami jelentősen javult a mintahossz duplázásával (várható volt a korábbiak alapján), ugyanakkor a többi módszernek is javult a teljesítménye.

A szakaszt összefoglalva azt mondhatjuk, hogy rövid minták esetében nem célszerű a Bernstein módszert alkalmazni, illetve általában – mivel a módszerek határeloszlási tételeken nyugszanak – érdemes hosszabb adat-sorokat bevinni az elemzésbe.

	β	HJ	PD	CB	CI	GR
DGP1	0.00	0.043	0.038	0.038	0.032	0.040
	0.10	0.107	0.129	0.110	0.152	0.743
	0.20	0.578	0.691	0.423	0.631	0.998
	0.30	0.960	0.988	0.906	0.981	1.000
	0.40	0.999	1.000	0.996	1.000	1.000
	0.50	1.000	1.000	1.000	1.000	1.000
DGP2	0.00	0.043	0.033	0.041	0.051	0.042
	0.10	0.282	0.306	0.246	0.081	0.090
	0.20	0.933	0.951	0.887	0.243	0.167
	0.30	1.000	1.000	1.000	0.610	0.249
	0.40	1.000	1.000	1.000	0.862	0.308
	0.50	1.000	1.000	1.000	0.984	0.362
DGP3	0.00	0.038	0.029	0.637	0.038	0.033
	0.10	0.067	0.054	0.713	0.150	0.034
	0.20	0.213	0.210	0.958	0.634	0.067
	0.30	0.617	0.640	0.999	0.975	0.112
	0.40	0.936	0.954	1.000	1.000	0.171
	0.50	0.999	0.999	1.000	1.000	0.244
DGP4	0.00	0.036	0.034	0.641	0.033	0.059
	0.10	0.920	0.881	0.923	0.063	0.095
	0.20	0.998	0.997	0.991	0.093	0.108
	0.30	1.000	1.000	0.998	0.120	0.140
	0.40	1.000	1.000	1.000	0.143	0.178
	0.50	1.000	1.000	1.000	0.155	0.160

4. táblázat. $T = 500$, $N = 1000$, $\beta \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$,
nominális szignifikancia szint = 5%

Irodalom

1. Baek, E.–Brock, W.: A general test for nonlinear Granger causality: Bivariate model, *Working paper*. Iowa State University and University of Wisconsin, Madison, 1992.
2. Baek, E.–Brock, W.: A nonparametric test for independence of a multivariate time series, *Statistica Sinica*. Vol. 31, No. 2 (1992), pp. 137–156.
3. Brown, Morton B.: 400: A Method for Combining Non-Independent, One-Sided Tests of Significance, *Biometrics* No. 4 (1975), pp. 987–992,

4. Csörgő Sándor–Faraway, Julian J.: The Exact and Asymptotic Distributions of Cramér–von Mises Statistics, *Journal of the Royal Statistical Society. Series B (Methodological)* Vol. 58, No. 1 (1996), pp. 221–234.
5. Diks, Cees–Panchenko, Valentyn: A new statistic and practical guidelines for nonparametric Granger causality testing, *Journal of Economic Dynamics & Control*. Vol. 30, No. 9–10 (2006), pp. 1647–69.
6. Fisher, R. A.: Answer to Question 14 on Combining independent tests of significance, *The American Statistician*. Vol. 2 (1948), p. 30.
7. Florens, J.-P.–Mouchart, M.: A Linear Theory for Noncausality, *Econometrica*. Vol. 53, No. 1 (1985), pp. 157–176.
8. Genest, Christian–Rémillard, Bruno: Test of independence and randomness based on the empirical copula process, *TEST*. Vol. 13, No. 2 (2004), pp. 335–369.
9. Genest, Christian–Queissy, Jean-Francois—Rémillard, Bruno: Asymptotic local efficiency of Cramer–von Mises tests for multivariate independence, *The Annals of Statistics*. Vol. 35, No. 1 (2007), pp. 166–191.
10. Geweke, John: Measurement of Linear Dependence and Feedback Between Multiple Time Series, *Journal of the American Statistical Association*. Vol. 77, No. 378 (1982), pp. 304–313.
11. Granger, Clive W. J.: Investigating causal relations by econometric models and cross-spectral methods, *Econometrica*. Vol. 37, No. 3 (1969), pp. 424–438.
12. Granger, Clive W. J.: Economic Processes Involving Feedback, *Information and Control* Vol. 6, No. 1 (1963), pp. 28–48.
13. Granger, Clive W. J.–Morgenstern, Oskar: Spectral Analysis of New York Stock Market Prices, *Kyklos*. Vol. 16, No. 1 (1963), pp. 1–27.
14. Granger, Clive W. J.–Newbold, Paul: *Forecasting Economic Time Series*. Academic Press (1977).
15. Hiemstra, Craig–Jones, Jonathan D.: Testing for Linear and Nonlinear Granger Causality in the Stock Price-Volume Relation, *Journal of Finance*. Vol. 49, No. 5 (1994), pp. 1639–64.
16. Hofert, Marius–Kojadinovic, Ivan–Maechler, Martin–Yan, Jun: copula: Multivariate Dependence with Copulas. *R package* version 0.999-13 (2015) <http://CRAN.R-project.org/package=copula>
17. Karpoff, Jonathan: The Relation Between Price Changes and Trading Volume: A Survey, *The Journal of Financial and Quantitative Analysis*. Vol. 22, No. 1 (1987), pp. 109–126.
18. Poole, William–Gibbs, David L.–Shmulevich, Ilya–Bernard, Brady–Knijnenburg, Theo: Combining Dependent P-values with an Empirical Adaptation of Brown’s Method, DOI: <http://dx.doi.org/10.1101/029637>
19. Sancetta, Alessio–Satchell, Stephen: The Bernstein Copula and Its Applications to Modeling and Approximations of Multivariate Distributions, *Econometric Theory*. Vol. 20, No. 3 (Jun., 2004), pp. 535–562.
20. Sklar, A.: Fonctions de répartition à n dimensions et leurs marges, *Publ. Inst. Statist. Univ.* Vol. 8, pp. 229–231.
21. Su, Liangjun–White, Halbert: A consistent characteristic function-based test for conditional independence, *Journal of Econometrics*. Vol. 141, No. 2 (Dec., 2007), pp., 807–834.

22. Sun, Xiaohai: Assessing Nonlinear Granger Causality from Multivariate Time Series, In *Machine Learning and Knowledge Discovery in Databases*. Springer Berlin Heidelberg (2008), pp. 440–455.
23. Bouezmarni, Taoufik–Rombouts, Jeroen V.K.–Taamouti, Abderrahim: A Non-parametric Copula Based Test for Conditional Independence with Applications to Granger Causality, *Working Paper, Universidad Carlos III de Madrid*. Vol 19 (2009)
24. Taamoutia, Abderrahim–Bouezmarni, Taoufik–El Ghouhc, Anouar: Non-parametric estimation and inference for conditional density based Granger causality measures, *Journal of Econometrics*. Vol 180, No. 2 (2014), pp. 251–264.
25. Varga József: Kopulák alkalmazása a pénzügyi kockázat menedzsmentben: Matematikai alapok, *SZIGMA*. 35. évf. 3-4. sz. (2004), 91–106. o.
26. Wiener, Norbert: *The theory of prediction*, Modern Mathematics for Engineers. McGraw-Hill (1956).
27. Zeileis, Achim–Hothorn, Torsten: Diagnostic Checking in Regression Relationships. *R News* Vol. 2, No. 3 (2002), pp. 7–10. <http://CRAN.R-project.org/doc/Rnews/>

NON-PARAMETRICAL CAUSALITY TESTS FOR TWO VARIABLES

In this paper we present the most important non-parametric alternatives of the classical Granger-causality furthermore introduce a new method based on independent copulas. After describing the causality tests we compare them simulating different non-linear data generating processes. The results show that our new method performs better in some cases. The paper concludes by analysing a well-known empirical problem in finance: the relation between return and trading volume of the Dow Jones Index.