

DOKTORI DISSZERTÁCIÓ

**Pénzügyi folyamatok elemzése és  
modellezése statisztikus fizikai módszerekkel**

**Kullmann László**

Témavezető: PROF. KERTÉSZ JÁNOS  
MTA LEVELEZŐ TAGJA

*Budapesti Műszaki és Gazdaságtudományi Egyetem  
Elméleti Fizika Tanszék*

2003

# Tartalomjegyzék

<b>1. Bevezetés</b>	<b>1</b>
<b>2. Irodalmi áttekintés</b>	<b>4</b>
2.1. Tőzsdeindex empirikus vizsgálata . . . . .	4
2.1.1. Index fluktuációjának eloszlása . . . . .	6
2.1.2. A hozam és abszolút értékének korrelációja . . . . .	11
2.1.3. Hozam eloszlásának skálázása . . . . .	16
2.1.4. ARCH, GARCH modellek . . . . .	19
2.2. Keresztkorrelációk részvényárfolyamok közt . . . . .	22
2.2.1. Csoportosítási módszerek . . . . .	25
2.3. Tőzsdei sok ügynökös modellek . . . . .	30
2.3.1. „Csorda”-hatás . . . . .	30
2.3.2. Aktivitás . . . . .	33
<b>3. Tőzsdeindex fluktuációjának vizsgálata</b>	<b>36</b>
3.1. Eloszlás skálázásának határa . . . . .	36
3.1.1. Illesztési eljárás . . . . .	39
3.1.2. Karakterisztikus idők meghatározása . . . . .	41
3.2. Egyrészvényes mikroszkopikus modell . . . . .	45
3.2.1. Csoportok kialakításának modellje . . . . .	45
3.2.2. A csoportméret eloszlás teljes megoldása . . . . .	47
3.2.3. Nemlineáris aktivitás . . . . .	48
3.2.4. Eredmények . . . . .	50
3.2.5. Átmenet a Gauss eloszlásba a „csordaszellemű” modellek esetén . . . . .	51
<b>4. Részvény árfolyamok keresztkorrelációi</b>	<b>56</b>
4.1. Időfüggetlen keresztkorrelációk vizsgálata . . . . .	56
4.1.1. Szuper-paramágneses Potts modell alkalmazása . . . . .	57
4.1.2. Eredmények értékelése . . . . .	59
4.2. Időfüggő keresztkorrelációk vizsgálata . . . . .	63
4.2.1. Felhasznált adatok . . . . .	64

4.2.2. Zajredukció . . . . .	65
4.2.3. Eredmények . . . . .	68
<b>Összefoglalás</b>	<b>72</b>
<b>Köszönetnyilvánítás</b>	<b>75</b>
<b>Függelék</b>	<b>75</b>
<b>A. Csoportméret eloszlások a mikroszkopikus modellben</b>	<b>76</b>
A.1. Csoportméret aszimptotikus eloszlása . . . . .	77
A.2. Az egyes csoportok méreteloszlásának analitikus megoldása . . . . .	78
A.3. Csoportok átlagos mérete . . . . .	81
A.4. A csoportméret eloszlás, $\mathbf{P}(s, t)$ , időfüggő megoldása . . . . .	81
A.5. Aszimptotikus viselkedések . . . . .	82
<b>B. Valószínűségi változók összegének szórásnégyzete és a változók autokorrelációjának exponense közti összefüggés levezetése</b>	<b>85</b>
<b>C. Vállalatok neve, szimbóluma, és rövid leírása</b>	<b>87</b>
<b>Irodalomjegyzék</b>	<b>88</b>
<b>Publikációs lista</b>	<b>95</b>

# 1. fejezet

## Bevezetés

Az utóbbi időben a statisztikus fizika alkalmazási köre túllép a hagyományos vizsgálódási területeken. Ehhez jelentősen hozzájárult a kaotikus rendszerek viselkedésének megértése, a fraktálszerkezet felfedezése, a sztochasztikus folyamatok leírása, illetve a renormálási csoport transzformáció kidolgozása. Ezen fogalmak illetve módszerek mindenhol alkalmazhatók, ahol sok szabadsági fokú, kölcsönható rendszerek leírása a feladat, mint például a biológiában, a szociológiában, de jól használhatók autópályamodellek felállításában, vagy a granuláris anyagok áramlásának vizsgálatánál, és a legújabb eredmények azt igazolják, hogy a gazdasági folyamatok elemzésénél is. Az utóbbi alkalmazási terület önálló nevet is kapott: gazdaság-fizika (econophysics).

A gazdaság-fizika egyik legfontosabb területe az árfluktuációk vizsgálata. A gyakorlati alkalmazás szempontjából (pl. opciók árazása) is rendkívül fontos az árváltozások empirikus adatokon megfigyelt viselkedésének értelmezése és modellezése.

A XX. század elején Bachelier a tőzsdei ár időbeli fluktuációját véletlen bolyongási modellel kísérelte meg leírni [Bachelier, 1900]. Az árváltozások idősorát vizsgálva azonban azt tapasztalhatjuk, hogy a nagy ugrások jóval gyakoribbak, mint egy Gauss-eloszlással jellemzett véletlen bolyongás esetén. Ezen megfontolásból alkalmasabbnak látszott az áringadozásokat hatványfüggvény eloszlással rendelkező Lévy folyamatokkal leírni [Mandelbrot, 1963]. Az újabban rendelkezésre álló, ún. nagy felbontású adatok elemzése megerősítette azt a megállapítást, hogy az árfluktuációk normálisnál lassabban lecsengő eloszlással rendelkeznek [Mantegna and Stanley, 1995].

A hatványfüggvény-viselkedést először fizikai rendszerekben a másodfajú fázisátalakulásoknál sikerült megérteni. Mint ismeretes, a kritikus pont közelében a rendszer – mivel a karakterisztikus méret, a korrelációs hossz divergál –, önhasnó tulajdonsággal rendelkezik, skálainvariáns. A skálázást a gazdasági folyamatoknál is sikerült kimutatni. A különböző időkülönbségekhez tartozó növekmények (árváltozás) sűrűségfüggvényei nem triviális módon egymásba skálázhatók [Mantegna and Stanley, 1995]. Ennek az oka abban rejlik, hogy az hatványfüggvény lecsengésű eloszlás exponense,  $\alpha < 2$ , tehát a megfelelő Lévy eloszlás stabil. Ugyanakkor nagyfrekven-

ciás adatokon (a tranzakciók gyakorisága másodpercnyi nagyságrendű) történt vizsgálatok azt mutatták, hogy az árfluktuáció eloszlásában megfigyelhető egy levágás [Gopikrishnan et al., 1999], – nagy árugrások esetén az eloszlás lecsengési exponense,  $\alpha > 2$  lesz –, ami azt eredményezi, hogy a skálázás nem minden határon túl folytatható, és a nagy időkülönbségekhez tartozó árváltozásoknál normális eloszlás várható. Fontos kérdés tehát, hogy milyen karakterisztikus idők lépnek fel a rendszerben. Az eloszlások pontosabb megismerése bővíti az ú.n. stilizált tények körét, vagyis azon megfigyelések csoportját, amely a tőzsdei adatokat jellemzi és amelyeket a „mikroszkopikus”, sokügynökös modelleknek vissza kell adni.

A portfólió-optimalizálás szempontjából, de önmagában is fontos kérdés a különböző piaci egységek (részvény, egyes tőzsdék) közti kölcsönhatások tanulmányozása. Köztudott, hogy a különböző részvények árfolyamai egymástól nem függetlenek. Hasonló kapcsolat figyelhető meg különböző tőzsdék indexértékeinek időfejlődésében is.

A kölcsönhatások vizsgálatának legalkalmasabb módszere a különböző részvények (vagy tőzsdék) árváltozásai közt mért korrelációk elemzése. Az egyidejű korrelációk – két különböző részvény ugyanabban az időpontban bekövetkezett árváltozásának korrelációja – vizsgálata alapján a részvények jól elkülönülő csoportokba rendezhetők [Mantegna, 1999]. Azonos csoportba azon részvények tartoznak, amelyek közt a korreláció erőssége nagyobb, mint egy meghatározott küszöbérték. Ezen küszöbérték változtatásával más-más csoportok fognak kialakulni. Csökkentve a küszöbértéket, az egyes csoportok alcsoportokra esnek szét, majd tovább csökkentve a küszöböt ezek újabb alcsoportokra bomlanak, ezáltal hierarchikus struktúra alakul ki. A portfólió-optimalizáláshoz a részvények csoportosításánál különös jelentőséggel bírhat a negatív korrelációk figyelembe vétele.

A részvények közti időfüggő korreláció vizsgálata további információval szolgál a részvények közti kölcsönhatások megértéséhez. Amennyiben a részvényárfolyamok közti idő-eltolás függvényében mért korreláció aszimmetrikus (tehát a korreláció maximuma nem az egyidejű árugrásoknál mutatkozik), azt jelenti, hogy az egyik részvélynél hamarabb következett be az ugrás, tehát ez a részvény úgy mond „vonzza” a másikat. Ha minden részvénypárra megvizsgáljuk ezt az összefüggést, egy irányított hálózatot kapunk. Ez nem csak azt mutatja meg, hogy mely részvények tartoznak egy csoportba, hanem a csoportot domináló részvényekről is információval szolgál, tehát arról, hogy melyik az a részvény, amelyik változása a többit befolyásolja.

A disszertációban az itt említett kérdésekkel kapcsolatos eredményeket mutatunk be. A dolgozat felépítése a következő:

A 2. fejezet egy irodalmi áttekintést nyújt. Ennek a szokásosnál kissé bővebb terjedelmét az indokolja, hogy a gazdaság-fizika még fiatal tudományág, amelynek eredményei és vizsgálódási területei még nem ismertek széles körben. Ezért fontosnak tartottuk, hogy leírjuk az alapvető fogalmakat, illetve a fő kérdésköröket.

A dolgozat ezt követően két nagyobb témakört érint. Az első (3. fejezet) az

árfluktuációk vizsgálatával foglalkozik.

Elsőként a 3.1 fejezetben a tőzsdeindex fluktuációjának eloszlását tanulmányozzuk. A különböző időkülönbségekhez tartozó indexfluktuációk empirikus eloszlásfüggvényeire normális, illetve Lévy eloszlású függvényeket illesztettünk. Célunk a különböző karakterisztikus idők meghatározása és leírása volt.

A 3.2 fejezetben egy mikroszkopikus modellel foglalkozunk, amely a valós részvényárfolyam időfejlődését hivatott modellezni. A modell alapja Cont „csordamodell”-je [Cont and Bouchaud, 2000], amin két változtatást hajtottunk végre. Egyrészt a csoportok kialakítása a Barabási féle hálózati modell [Albert et al., 1999] által inspirált módon történt. Másrészt az adott időpillanatban aktív ügynökök számát az árfolyam nemlineáris függvényének választottuk. Azt vizsgáltuk, hogy az így kialakított modell mennyire egyezik eredményeikben a valós árfluktuációk tulajdonságaival.

A másik nagyobb témakörben (4 fejezet) a részvények között kialakuló kölcsönhatásokat elemezzük a mért korrelációk segítségével.

A 4.1 fejezet témája a részvények csoportokba rendezése az időfüggetlen korrelációk felhasználásával. Ehhez egy mágneses modellt, a szuper-paramágneses klaszterezést alkalmaztunk. Ezt úgy módosítottunk, hogy ne csak ferromágneses, hanem anti-ferromágneses kölcsönhatásokat is figyelembe tudjunk venni. Ennek segítségével meghatároztuk mind a Dow Jones mind a S&P500 tőzsdeindex részvényei közt kialakuló csoportokat, és eredményeinket összehasonlítottuk a már korábban, más módszerrel meghatározott eredményekkel [Mantegna, 1999].

Végül a 4.2 fejezetben az időfüggő keresztkorrelációk vizsgálatával foglalkoztunk. Egy nagyfrekvenciás – az árváltozások időkülönbsége másodperc nagyságrendű – adathalmazon vizsgáltuk az időfüggő korrelációk szimmetria tulajdonságait. Noha az ebből adódó effektus kicsi, mégis van jelentősége: ennek segítségével a részvények között egy irányított befolyásolási hálózatot határoztunk meg.

## 2. fejezet

# Irodalmi áttekintés

### 2.1. Tőzsdeindex empirikus vizsgálata

A tőzsdeindex a tőzsdén szereplő legfontosabb részvények árfolyamaiból képzett súlyozott átlag. A 2.1/a) ábrán egy példa látható, amely a Standard and Poor's 500 index időbeli változását mutatja. A tőzsdeindex az egyik legfontosabb mennyiség, mellyel a tőzsdét jellemezni lehet, így nem meglepő, hogy ennek a tanulmányozása a gazdaság-fizika egyik központi kérdése. Az index értékének időbeli változása és az egyedi részvények árfluktuációja közt szoros párhuzam vonható, így a továbbiakban csak általánosan az árváltozásról beszélünk.

A részvény árát általában egy ország valutanemében adják meg. Az előbbi azonban nem időfüggetlen, hiszen infláció, gazdasági növekedés, vagy a valutanem értékének fluktuációja is megváltoztathatja értékét. Emiatt nem mindegy, hogy milyen módon definiáljuk a részvény értékének változását. Jelölje  $p(t)$  az ár értékét a  $t$ -ik időpillanatban. Ekkor a  $\Delta t$  időkülönbséghez tartozó árváltozásra a leggyakrabban használt mennyiségek:

növekmény:

$$z_{\Delta t}(t) = p(t + \Delta t) - p(t) \quad (2.1)$$

hozam:

$$s_{\Delta t}(t) = \frac{p(t + \Delta t) - p(t)}{p(t)} = \frac{z_{\Delta t}(t)}{p(t)} \quad (2.2)$$

logaritmikus hozam:

$$r_{\Delta t}(t) = \log [p(t + \Delta t)] - \log [p(t)] \quad (2.3)$$

diszkontált növekmény:

$$z_{D,\Delta t}(t) = [p(t + \Delta t) - p(t)]D(t), \quad (2.4)$$

ahol  $D(t)$  a diszkontálási faktor, ami a banki kamat segítségével veszi figyelembe a fent említett hatásokat.

A gazdaság-fizikai tárgyalásmódban általában a logaritmikus hozam definíciót használják, ennek okát később részletezzük. Nagy frekvenciás adatok esetén, amikor  $\Delta t$  kicsi, fennáll, hogy  $|z_{\Delta t}(t)| \ll p(t)$ , ezért  $s_{\Delta t}(t) \simeq r_{\Delta t}(t)$ , azaz a hozam és a logaritmikus hozam egyaránt használható.

Mint láthatjuk, az ár változásának egyik fontos paramétere, hogy mekkora idő alatt következik be az adott változás. Az idő definiálása a gazdasági folyamatokban éppen olyan nehéz, mint az árváltozásé, hiszen a tranzakciók, amelyek meghatározzák az ár értékének kialakulását, időben véletlenszerűen követik egymást. Nyilvánvaló az is, hogy csak a tőzsde nyitvatartása alatt történhetnek. Ebből kifolyólag itt is többféle definíció vezethető be. (i) *Fizikai idő*, ami azt jelenti, hogy nincs kitüntetve az az idő, amikor a tőzsde nyitva van, tehát ugyanakkora súllyal számítja az éjjeli, illetve hétvégi időfolyamatot. (ii) *Üzletelési idő*, csak azt az időt veszi figyelembe, amikor a tőzsde nyitva van. Ennek az a hátránya, hogy noha éjjel nincs üzletelés, mégis információhoz jutnak az emberek, amely alapján reggel döntéseket hoznak, az üzletelés szempontjából ezért ez az idő nem indifferens. Továbbá a számítógépek lehetővé tesznek nyitvatartáson kívüli tranzakciót. Mindkét idődefiníció további hibája, hogy az aktivitás nem egyenletes, azaz két üzletkötés közti időintervallum is változik a nap folyamán. (iii) Ez utóbbi miatt született a *tranzakciós idő*, ami két üzletkötés közt eltelt időt vesz egységnek [Zumbach, 1998].

Az ár jövőbeli értékének meghatározására nincs determinisztikus modell. Ennek oka nem csak abban rejlik, hogy a befolyásoló tényezők száma nagy, és így gyakorlatilag nem áll elég információ a rendelkezésre, hanem a determinisztikus meghatározás még elvben sem lehetséges. Egy ilyen modell ismeretében lehetséges lenne pusztán spekuláció útján kockázatmentes nyereséghez jutni, úgynevezett *arbitrage* üzleteket kötni. Természetesen ezt mindenki a saját hasznára szeretné fordítani, és ezek a spekulációs üzletek a modelltől eltérítenék az ár alakulását [Cont et al., 1997, Mantegna and Stanley, 2000]. Az olyan piacokat, ahol nem tud kialakulni arbitrage *hatékony piacoknak* nevezzük, azaz a piac akkor hatékony, ha a piacot befolyásoló minden információ azonnal tükröződik az ár új értékében [Fama, 1970]. Természetesen a teljesen hatékony piac csak idealizált rendszer, a valós piacok ettől eltérőek.

Az empirikus vizsgálatok alapján a tőzsdei logaritmikus hozam olyan sztochasztikus folyamatnak felelnek meg, amik az alábbi tulajdonságokkal jellemezhetők:

*i)* valószínűségrészsűrűség függvényük hatványfüggvény lecsengést mutat  $\alpha \simeq 4$  körüli exponenssel [Gopikrishnan et al., 1999]

*ii)* nagyobb (több hónapos) időkülönbséget tekintve a sűrűségfüggvény Gauss eloszláshoz konvergál [Gopikrishnan et al., 1999, Kullmann et al., 1999, Kullmann et al., 2000b]

*iii)* gyengén korreláltak, a korreláció karakterisztikus értéke a néhány perc tarto-

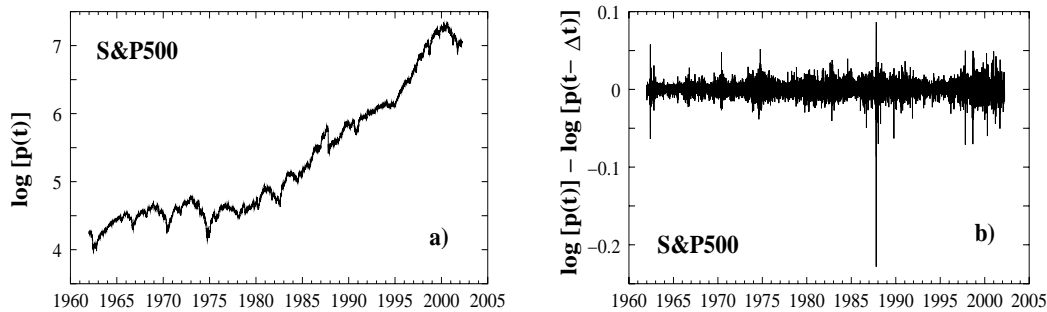
mányába esik [Liu et al., 1999, Cont et al., 1997]

*iv)* abszolút értékük korrelációja lassú, hatványfüggvény-szerű lecsengést mutat [Gopikrishnan et al., 1999]

*v)* időfejlődésük időben aszimmetrikus [Arnéodo et al., 1998]

*vi)* a hozam  $q$ -ik momentumának autokorrelációja multiskálázást mutat, azaz a hatványfüggvény lecsengésű autokorreláció exponense  $q$ -nak nemlineáris függvénye. [Pasquini and Serva, 2000]

A továbbiakban ezen tulajdonságokat röviden elemezzük.



2.1. ábra. Az S&P 500 tőzsdeindex időbeli változása. Az **a)** ábrán az index érték logaritmusának időbeli fejlődése látható. A **b)** ábrán a logaritmusos árkülönbség (hozam) időfejlődése látható.

### 2.1.1. Index fluktuációjának eloszlása

A tőzsdeindex fluktuációjának leírására már a XX. század elején készültek modellek [Bachelier, 1900]. Bachelier az index fluktuációit véletlen bolyongással, nevezetesen a Brown mozgással kísérelte meg leírni. A feltételezés magyarázata abban keresendő, hogy a természetben megfigyelt véletlen folyamatok gyakran mutatnak normális eloszlást. Ennek gyökere a centrális határeloszlás tételében rejlik, amire a dolgozat a későbbiekben fog részletesebben kitérni.

Később, a 60-as években Mandelbrot azt figyelte meg, hogy az index fluktuációjában a nagy ugrások sokkal nagyobb valószínűséggel fordulnak elő, mint amit a normális eloszlás jósolna, azaz a valószínűsége függvénynek lassabban kell lecsengenie [Mandelbrot, 1963]. A nagy árugrások gyakoriságának szemléltetésére jó példa a 2.1 ábra. Egy Gauss eloszlásnak megfelelő véletlen bolyongás esetén gyakorlatilag sosem kapnánk akkora ugrásokat, mint például az 1986-ban bekövetkező „fekete hétfő” esetén.

## Lévy eloszlás

Mandelbrot az index fluktuációk eloszlását a Lévy eloszlással jellemezte, amelynek karakterisztikus függvénye az alábbi alakot ölti:

$$f_\alpha(z) = \exp \left\{ i\gamma z - a_\alpha |z|^\alpha \left( 1 + i\beta \tan(\alpha\pi/2) \frac{z}{|z|} \right) \right\}. \quad (2.5)$$

A kifejezésben szereplő  $\alpha$  exponens  $0 < \alpha < 2$  értékeket vehet fel. ( $\alpha = 2$  esetén a normális eloszlást kapjuk.) A  $\beta$  paraméter az eloszlás ferdeségét adja meg,  $\beta = 0$  felel meg a szimmetrikus eloszlásnak, illetve a  $\gamma$  várhatóérték. A tapasztalat azt mutatta [Lillo and Mantegna, 2000b], hogy a drift effektusok okozta árváltozás elhanyagolható az átlag körüli fluktuációkhoz – más néven *volatilitáshoz* – képest, ezért az árfluktuációk vizsgálatánál leggyakrabban szimmetrikus,  $\beta = 0$  eloszlást használnak. Mivel továbbá feltételezzük, hogy a piac hatékony (nincs arbitrage), az eloszlásfüggvény várhatóértéke jó közelítéssel nulla,  $\gamma = 0$ , tehát a logaritmikus hozam sűrűségfüggvénye legjobban az alábbi alakkal jellemezhető:

$$f(z) = \exp\{-a_\alpha |q|^\alpha\}. \quad (2.6)$$

A sűrűségfüggvény szimmetriájára tett állítás nem teljesül olyan esetekben, amikor a tőzsdén nagy árzuhanás vagy nagy áremelkedés következik be. Az ilyen napokon, illetve ezen napok környékén az eloszlás erősen aszimmetrikus lesz [Lillo and Mantegna, 2000a]. A szerzők azt vizsgálták, hogy különböző részvények adott napon vett hozamaiból alkotott sokaság eloszlása hogyan változik az időben. Azt tapasztalták, hogy abban az esetben, amikor nincs extrém árváltozás (árugrás vagy áresés), a sokaság eloszlása szimmetrikus. Ugyanakkor extrém árváltozások esetén az eloszlás alakja megváltozik: árzuhanás (áremelkedés) esetén az eloszlásfüggvényben negatív (pozitív) értékeknél egy csúcs jelenik meg, illetve a negatív (pozitív) oldal meredeksége nagyobb lesz.

A Lévy eloszlás sűrűségfüggvényének – amely a karakterisztikus függvény inverz Fourier transzformáltja –, nincs analitikus alakja. Ugyanakkor nagy valószínűségi változók esetén sorfejtést lehet alkalmazni, amely az alábbi, hatványfüggvény lecsengésű alakban írható:

$$L_\alpha(|x| \rightarrow \infty) \sim \frac{\alpha A_\pm^\alpha}{|x|^{1+\alpha}}. \quad (2.7)$$

A Lévy eloszlás stabil, ami azt jelenti, hogy a független, Lévy eloszlású valószínűségi változók összege szintén Lévy eloszlású lesz.

Az eloszlás normálistól való eltéréseinek tanulmányozására többféle módszer is kínálkozik, ezek közül az egyik a *kurtózis* vizsgálata, amely egy valószínűségi változó negyedik normált félinvariánsából származtatható:

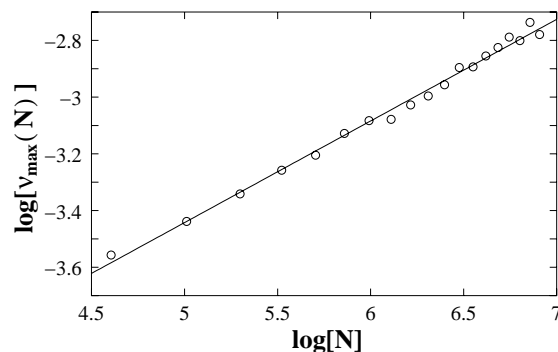
$$\kappa = \frac{\langle (x - \langle x \rangle)^4 \rangle}{\sigma^4}, \quad (2.8)$$

ahol  $x$  a valószínűségi változó,  $\sigma$  a szórása, és a  $\langle \cdot \rangle$  jelölés a valószínűségi átlagot jelenti. Normális eloszlás esetén  $\kappa = 3$ . Az S&P 500 indexen végzett számítások a  $\Delta t = 5$  perchez tartozó hozam esetén  $\kappa = 19$ -et eredményeztek [Cont et al., 1997]. A pozitív kurtózis érték a Gauss eloszlásnál lassabban lecsengő eloszlásra utal.

Az eloszlás normális jellegének egy másik lehetséges ellenőrzése az eloszlás extrém értékeinek vizsgálata [Bouchaud and Potters, 1997]. Jelölje  $\nu_{max}$  az ismert eloszlású valószínűségi változók  $N$  elemű mintájából vett maximális érték várhatóértékét. Ez a változó nyilván függ az eloszlástól, illetve  $N$ -től. Az eljárást vizsgáljuk meg két példán keresztül, amelyek lényegesen eltérő eredményeket szolgáltatnak.

1.  $P(x) \sim e^{-\beta x}$ , ha  $x \rightarrow \infty$ . Ekkor:  $\nu_{max}(N) = \frac{\log(\frac{N}{\beta})}{\beta}$
2.  $P(x) \sim \frac{A_+^\alpha}{x^{1+\alpha}}$ , ha  $x \rightarrow \infty$ . Ekkor:  $\nu_{max}(N) = A_+ N^{1/\alpha}$

A Standard & Poor's 500 indexen végzett analízis azt mutatja, hogy az  $N$  elemű minta maximális értékének várhatóértéke a minta elemszámának hatványfüggvénye (2.2 ábra), ami alátámasztja, hogy a hozam értékek eloszlása hatványfüggvény jelleggel cseng le.



2.2. ábra. Az S&P500 index egy napos időkülönbséghez tartozó hozam értékeinek extrém-érték statisztikája. Az vízszintes tengelyen a hozam értékekből vett minta elemszámának logaritmus, a függőleges tengelyen a megfelelő elemszámhoz tartozó extrém érték várható értékének logaritmus látható. Az illesztett egyenes meredekségének reciproka alapján a lecsengés exponensét adja:  $\alpha = 2.7896$

### Levágott Lévy eloszlás

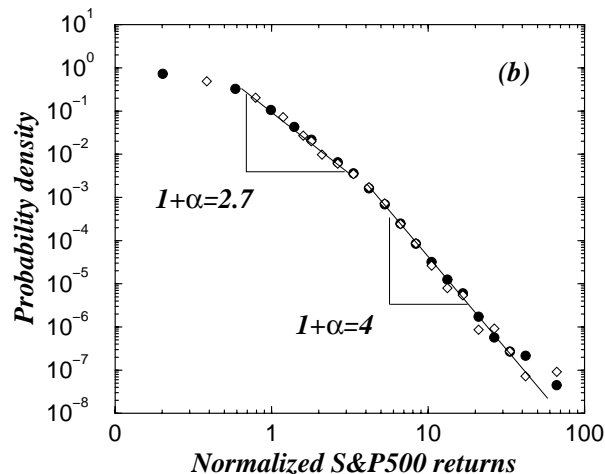
További vizsgálatok [Mantegna and Stanley, 1995, Cont et al., 1997] azt mutatták, hogy noha az empirikus adatok eloszlása lassabb lecsengést mutat a normális eloszláshoz képest, a Lévy eloszlás csak az origó közelében – amennyiben az empirikus eloszlás szórását  $\sigma$ -val jelöljük, akkor  $|r|/\sigma \lesssim 6$  esetén – mutat jó illeszkedést. A nagyobb árváltozások valószínűsége kisebb, mint amit egy Lévy eloszlás (2.5) ad, azaz az empirikus eloszlás leginkább egy „levágott” Lévy eloszlásnak felel meg.

Az árfluktuációk jellemzésére eleinte exponenciális levágású Lévy eloszlást javasoltak [Mantegna and Stanley, 1995, Cont et al., 1997], amely a következő karakterisztikus függvénnyel jellemezhető:

$$f_{\alpha,\mu}(z) = \exp \left\{ -a_{\alpha} \frac{(\mu^2 + z^2)^{\alpha/2} \cos(\alpha \arctan(|z|/\mu)) - \mu^{\alpha}}{\cos(\pi\alpha/2)} \right\} \quad (2.9)$$

A kifejezésben szereplő  $\mu$  paraméter a levágási paraméter. Az egyenletből látszik, hogy  $\mu = 0$  esetén a (2.9) kifejezés megegyezik a (2.5) egyenlettel (ha  $\beta = 0$ ).

További, nagyobb mennyiségű adatokon végzett vizsgálatok azt mutatták [Gopikrishnan et al., 1998, Gopikrishnan et al., 1999], hogy a levágás is hatványfüggvény, de ennek az exponense  $\alpha \simeq 3$ , azaz kívül esik a stabil Lévy tartományon, és ezért vonatkozik rá a centrális határeloszlás tétel.



2.3. ábra. Standard & Poor's 500 tőzszeindex hozam értékeinek sűrűségfüggvénye [Gopikrishnan et al., 1999]. A vízszintes tengelyen a normált (a teljes vizsgált időtartomány szórással leosztott) hozam értékek, a függőleges tengelyen a hozzá tartozó valószínűségek szerepelnek. Kis hozam értékeknél az eloszlást stabil Lévy tartományba eső,  $\alpha = 1.7$ -es exponensű hatványfüggvénnyel, míg nagyobb hozam értékeknél a Lévy tartományon kívüli,  $\alpha \simeq 3$  exponenssel lecsengő hatványfüggvénnyel lehet jellemezni.

Az eloszlásfüggvény lecsengési exponensének mérésére több módszer is létezik. A legkézenfekvőbb módszer az eloszlás- vagy a sűrűségfüggvényre való egyenes illesztése log-log skálán. Egy másik eljárás az úgynevezett *Hill becslés* [Hill, 1975], amely az eloszlásfüggvény lokális meredekségének mérésén alapszik. Tegyük fel, hogy az eloszlás hatványfüggvény lecsengésű:  $P(\xi \geq x) = 1 - F(x) = C/x^{\alpha}$ , ekkor nyilván:

$$\frac{1}{\alpha} = - \left( \frac{d \log x}{d \log P(\xi \geq x)} \right) \quad (2.10)$$

Ha egy rendezett,  $N$  elemű mintát veszünk, azaz  $X_1 \geq X_2 \geq \dots X_N$ , ahol  $X_i$  jelenti az  $i$ -edik esemény (hozam) értékét a rendezett sorban, akkor a (2.10) egyenletet a

lokális meredekségekre átírva és felhasználva, hogy az így kapott empirikus eloszlásra  $P_N(\xi \geq X_k) = k/N$ :

$$\frac{1}{\alpha} = \frac{\log(X_i) - \log(X_{i+1})}{\log[P(X_{i+1})] - \log[P(X_i)]} = \frac{\log(X_i/X_{i+1})}{\log[(i+1)/i]} \quad (2.11)$$

A fenti kifejezést átlagolva az exponens Hill becslésére az alábbi egyenlet adódik:

$$\frac{1}{k} \sum_{i=1}^k \log \frac{X_i}{X_{k+1}} = H_{k,N} \xrightarrow[k, N \rightarrow \infty]{} \frac{1}{\alpha} \quad (2.12)$$

Ezek a módszerek azonban kis adatszám esetén könnyen adhatnak rossz exponens értéket, ami az általunk vizsgált esetben fontos jelentőséggel bír, hiszen bennünket az eloszlásfüggvény lecsengése érdekel, ahol a statisztika a leggyengébb. Az exponensérték rossz becslésének szemléltetésére jó példa az  $\alpha \lesssim 2$  exponenssel rendelkező Lévy eloszlású valószínűségi változók eloszlásfüggvényének vizsgálata. Azt tapasztalták [Weron, 2001], hogy amennyiben a generált valószínűségi változók száma kisebb mint  $10^8$ , a fent említett exponens becslési módszerek rossz, 2-nél nagyobb értékeket adnak.

### Opció ár számítása

Az *opció* olyan biztosítás, amely megvédi tulajdonosát a részvény árának nagy zuhanásától vagy ugrásától. Egy hosszú lejáratú üzletnél, a lejáratú idő elteltével az üzletfél, aki az opciót megvásárolta, szabadon dönthet arról, hogy teljesíti-e üzleti szándékát. Egy *Call* típusú opciónál (vételi jog vásárlása) például, ha a lejáratú idő után az előre lerögzített kötési árhoz képest csökken az áru értéke, az üzletfél lemond a vásárlási szándékáról, és így csak az opciós árat (a vételi jog ára) veszti el. Az arbitrage kizárásának érdekében az opció árának meg kell egyeznie az opciós üzlet kínálta nyereség várhatóértékével. Jelölje  $p_s$  a kötési árat, az értékpapír ára  $t = 0$ -ban legyen  $p_0$ , illetve  $T = N\tau$  idő múlva  $p(T)$  ( $\tau$  az elemi tranzakciók ideje). Ekkor az opciós ár [Potters et al., 1998]

$$C(p_0, p_s, T) = (1 + \rho)^{-N} \int_{p_s}^{\infty} (p' - p_s) P(p', T | p_0, 0) dp', \quad (2.13)$$

ahol  $\rho \equiv r\tau$  a diszkrét időkre számolt banki kamatláb, és  $P(p', T | p_0, 0) dp'$  annak a valószínűsége, hogy ha  $t = 0$  pillanatban az ár  $p_0$ -volt, akkor  $T$  idő elteltével  $p'$  értéket vesz fel. A nehézséget az adja, hogy  $P(p', T | p_0, 0)$  nem ismert.

Az opció árának meghatározására szolgál a Black-Scholes formula [Black and Scholes, 1973], ami azonban még normális eloszlású részvény árfolyamfluktuációt feltételez, és bemenő paraméterei az empirikus várható érték és szórás. Mivel

láttuk, hogy a valóságban az eloszlás nem normális, a Black-Scholes formula rossz opciós árat jelez előre, lehetőséget nyújtva ezáltal az arbitrage-nak.

Hogy ez ne valósuljon meg, az elmélet módosítása szükséges. Sokkal jobb becslés kapható [Potters et al., 1998], ha az árváltozások fluktuációjából,  $\sigma = \sqrt{\langle \delta x^2 \rangle}$  kapott érték helyett az alábbi tekintjük:

$$\Sigma(p_s, T) = \sigma \left[ 1 + \frac{\kappa_T}{24} \left( \frac{(p_s - p_0)^2}{\sigma_T^2} - 1 \right) \right] \quad (2.14)$$

Az egyenletben szereplő  $\sigma_T$  illetve  $\kappa_T$  mennyiségek a  $T$  időkülönbséghez tartozó szórás illetve kurtózis.

### 2.1.2. A hozam és abszolút értékének korrelációja

A tőzsdeindex empirikus vizsgálatának másik sokat analizált mennyisége az árváltozások (hozam) időbeli korrelációja.

$$C_{\Delta t}(t) = \frac{\langle r_{\Delta t}(t' + t) r_{\Delta t}(t') \rangle - \langle r_{\Delta t}(t' + t) \rangle \langle r_{\Delta t}(t') \rangle}{\sqrt{\langle (r_{\Delta t}(t') - \langle r_{\Delta t}(t') \rangle)^2 \rangle \langle (r_{\Delta t}(t' + t) - \langle r_{\Delta t}(t' + t) \rangle)^2 \rangle}}, \quad (2.15)$$

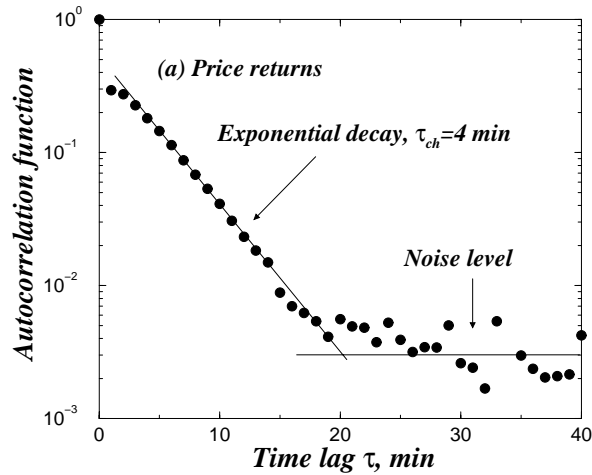
A kifejezésben szereplő  $\langle \cdot \rangle$  jelölés a teljes vizsgált időtartományra való átlagolást jelöli.

Nyilvánvaló, hogy amennyiben az időben egymást követő hozam értékek korrelálnának, azt jelentené, hogy az ár megjósolható lenne egy későbbi időpontban, ami viszont ellentmondana a 2.1 fejezetben definiált hatékony piac tulajdonságának. Ezt a gondolatmenetet igazolta a nagyfrekvenciás tőzsdeadatokon végzett analízis [Cont, 1997, Gopikrishnan et al., 1999], amely azt mutatta, hogy a  $\Delta t = 1$  perces időkülönbséghez tartozó hozamok korrelációja – (2.15) egyenlet –, exponenciálisan cseng le az idő függvényében. A karakterisztikus idő,  $\tau = \int_0^\infty C(t) dt$ , néhány perc csupán; ld 2.4 ábra.

Egy valószínűségi változó időbeli korrelációjának hiánya természetesen nem jelenti, hogy az időben egymást követő valószínűségi változók függetlenek lennének. A hozam értékek esetében is fennáll ez az állítás. Ha megnézzük a hozam abszolút értékének korrelációját, időben lassan, hatványfüggvény jelleggel lecsengő függvényt kapunk; ld. 2.5 ábra.

### Módszerek hosszútávú korrelációk mérésére

Az autokorreláció jellegének vizsgálatára több módszer létezik. Az első kézenfekvő, a korrelációs függvény időbeli viselkedésének tanulmányozása. Egy másik gyakran használt módszer a hatvány spektrálanalízis. A korrelációs függvény tulajdonképpen egy konvolúciós szorzat, aminek a Fourier transzformáltja egyszerűen a két idősor



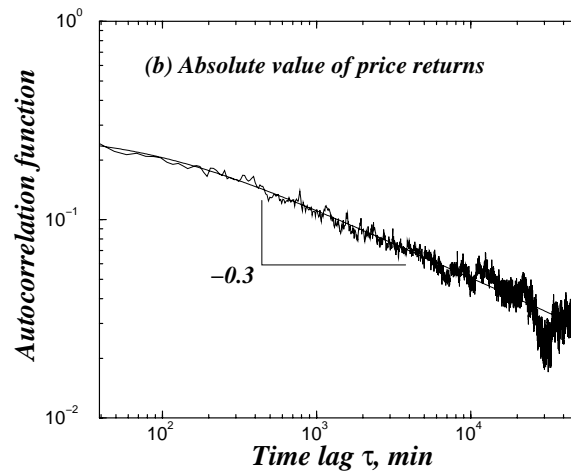
2.4. ábra. S&P500 tőzsdeindex hozam értékeinek autokorrelációja [Gopikrishnan et al., 1999]. A vízszintes tengelyen az idő látható lineáris skálán, percben mérve, a függőleges tengelyen a korrelációs függvény logaritmikus skálán. A korrelációs függvény exponenciális lecsengést mutat, karakterisztikus ideje 4 perc. Ez az érték összhangban van a hatékony piaccal.

Fourier transzformáltjának szorzata lesz, mivel a két idősor megegyezik (autokorreláció esetén). Végeredményben a korreláció Fourier transzformáltja az idősor Fourier transzformáltjának abszolút érték négyzete lesz (Wiener-Khinchin tétel):

$$\mathcal{F} \left[ \int_0^\infty f(t') f(t' - t) dt' \right] = |\mathcal{F}[f]|^2 \quad (2.16)$$

Fontos megjegyezni, hogy a hatvány spektrálanalízis csak stacionárius idősor esetén alkalmazható, ami azonban a tőzsdei folyamatoknál nem minden esetben érvényesül, mivel az árugrások szórása (volatilitás) időfüggő.

Elsősorban a hosszútávú, hatványfüggvény lecsengésű autokorrelációk vizsgálatára született a *Detrendált Fluktuáció Analízis* (DFA) módszere [Peng et al., 1994]. Ennek előnye, hogy nem egyensúlyi folyamatokban is lehetséges a hosszútávú korrelációk felismerése. A módszer a következő lépésekből áll: *i)* Az időbeli folyamatot  $l$  méretű, egyenlő nagyságú „dobozokra” osztjuk, majd minden dobozban egy egyenest illesztünk a legkisebb négyzetek módszerével, meghatározva ezáltal a lokális trendet a dobozon belül, (2.6 ábra). *ii)* Ezután meghatározzuk a detrendált idősort  $\tilde{y}(t)$ , ami az eredeti adatsor,  $y(t)$ , és a lokális trend különbsége. *iii)* Minden dobozban kiszámoljuk a detrendált adatsor szórását,  $\sigma_l^2 = \frac{1}{l} \sum_{t=1}^l \tilde{y}(t)^2$ , és ezeket átlagoljuk az összes  $l$  méretű dobozra,  $F^2(l) = \overline{\sigma_l^2}$ . A detrendált fluktuáció a doboz méretének hatványfüggvénye:  $F(l) \sim l^\kappa$ . Rövidtávú korrelációk esetén  $\kappa = 0.5$ , hosszútávú korrelációk esetén, azaz ha a korrelációs függvény hatványfüggvény lecsengést mutat:  $C(t) \sim t^{-\gamma}$ . A korrelációs exponens és a detrendált fluktuáció exponense közti összefüggés:  $\gamma = 2 - 2\kappa$ .



2.5. ábra. S&P500 tőzsdeindex hozam értékek abszolút értékének autokorrelációja log-log skálán [Gopikrishnan et al., 1999]. A vízszintes tengelyen az idő látható percben mérve, a függőleges tengelyen a korrelációs függvény. Az autokorreláció hatványfüggvény lecsengésű, a lecsengés exponense  $\gamma \simeq 0.3$ .

A hatvány spektrálanalízissel illetve a detrendált fluktuáció módszerével vizsgált volatilitás autokorreláció azt mutatta, hogy az autokorreláció – ellentétben a közvetlen, autokorrelációs függvény analízissel szemben – két tartományra szeparálódik az időben, ami két exponenssel jellemezhető:  $\gamma_1 = 0.14$  és  $\gamma_2 = 0.68$ . A két tartomány határához tartozó átcsapási idő  $t_x \simeq 600$  perc, ami körülbelül másfél napnak felel meg [Liu et al., 1999].

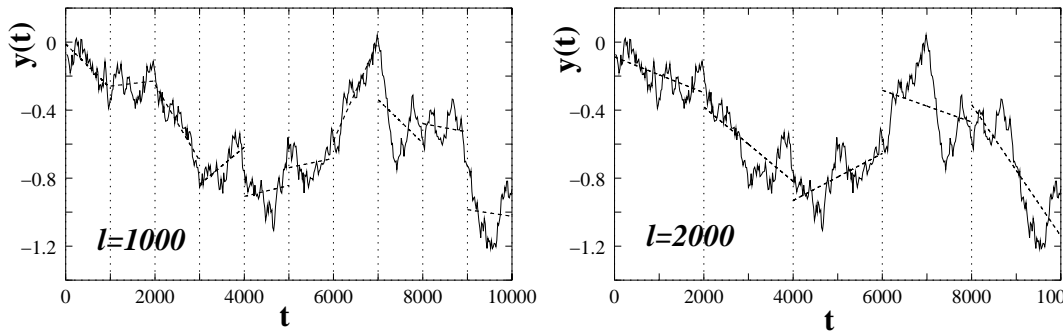
## Volatilitás

A gazdaságfizikában a kockázat mérésére gyakran használják a *volatilitás*-t, ami a tőzsdeindex, vagy egy részvény árának a várhatóértéke körüli fluktuációját írja le egy  $T = n\Delta t$  időtartományban.

$$v_T(t) = \frac{1}{n} \sum_{\tau=t}^{t+n-1} |r_{\Delta t}(\tau)| \quad (2.17)$$

Mivel a hozam értékének abszolút értéke (vagy négyzete) szorosan kapcsolódik a volatilitáshoz ezért az abszolút hozamra tett megállapítások általában igazak a volatilitásra is. Ezt a tényt igazolja, hogy a (2.17) egyenletben szereplő volatilitásnak mérték az autokorrelációját, és azt találták [Cizeau et al., 1997], hogy a  $\gamma \simeq 0.3$  exponensű lecsengés független attól, hogy mekkora  $T$  időtartományra átlagolom ki a hozam abszolútértékeket.

A volatilitás hosszútávú korrelációja a 2.1 ábrán is megmutatkozik. Ezen azt láthatjuk, hogy a nagy ugrások „fürtöket” alkotnak, azaz a nagy hozam értékek nem



2.6. ábra. Példa a DFA módszerre. A folytonos vonal egy generált véletlen bolyongás időfüggését mutatja. A szaggatott egyenesek a bolyongásnak a dobozon belüli lokális trendjét mutatják. Jól látszik, hogy nagyobb dobozméret esetén nagyobb lesz a lokális trend körüli szórás.

rendezetlenül helyezkednek el az időben, hanem egymáshoz közel, csoportokba tömörülve. Ezt a tulajdonságot a gazdaságfizikában a „volatilitás fürtöződésének” nevezik.

### Multifraktál tulajdonság

Mint utaltunk rá, a volatilitás a fluktuációkat méri. Definíciója azonban nem egyértelműen meghatározott, mivel vehetjük a hozam abszolút értékét, vagy négyzetét. Általános definícióként bevezethetjük az *általánosított abszolút hozam* fogalmát, ami egyszerűen a hozam abszolút értékének  $\nu$ -dik momentuma:  $|r_{\Delta t}(t)|^\nu$ . Ennek korrelációja:

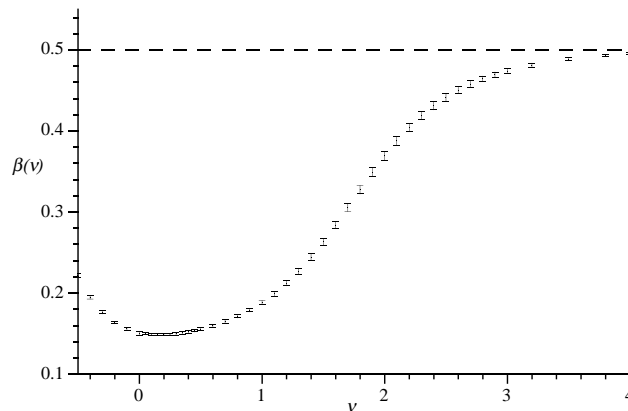
$$C_\nu(t) \sim \langle |r_{\Delta t}(t' + t)|^\nu |r_{\Delta t}(t')|^\nu \rangle - \langle |r_{\Delta t}(t' + t)|^\nu \rangle \langle |r_{\Delta t}(t')|^\nu \rangle, \quad (2.18)$$

szintén hatványfüggvény lecsengésű lesz  $C_\nu(t) \sim t^{-\gamma(\nu)}$ , de az exponens nemlineáris függvénye lesz a momentumnak,  $\nu$ -nek [Pasquini and Serva, 2000]. Ez a viselkedés olvasható le a 2.7 ábráról. Könnyen belátható ugyanis, hogy amennyiben az általánosított abszolút hozam korrelációja hatványfüggvény lecsengésű, egynél kisebb exponenssel, akkor a kumulált általánosított abszolút hozam,  $\chi(L, \nu) = \frac{1}{L} \sum_{i=1}^L |r_{\Delta t}(t + i)|^\nu$  szórása,  $\delta(L, \nu)$  is hatványfüggvény,  $\beta = \gamma/2$  lecsengéssel. A 2.7 ábrán a  $\beta$  exponens  $\nu$  momentumtól való függése látható.

A multifraktalitás abban nyilvánul meg, hogy a különböző  $\nu$  momentumok különböző tipikus fluktuációs méretet hoznak be a korrelációs átlagba, amik dominálni fogják a korrelációt. A  $\nu \gtrsim 4$  tartományban az átlagokat néhány, nagy hozam értékhez tartozó esemény dominálja, ahol a statisztika hatékonysága erősen lecsökken.

### Analógia a klasszikus diffúzióval

Ebben a fejezetben, valamint az előzőben azt láttuk, hogy a hozam eloszlása hatványfüggvény lecsengésű; időbeli korrelációja gyorsan, exponenciális függvénnyel cseng le néhány perces karakterisztikus idővel; míg abszolút értékének korrelációja lassú,



2.7. ábra. A kumulált általánosított abszolút hozam szórásának,  $\delta(L, \nu)$ -nek az exponense, mint a hozam momentumának,  $\nu$  függvénye, [Pasquini and Serva, 2000].

hatványfüggvény lecsengést mutat. Felmerülhet a kérdés, hogy mi okozza az eloszlásfüggvény hatványfüggvény lecsengését illetve a lassú lecsengésű korrelációját a hozam abszolút értékének. Ennek megértése a legegyszerűbb egy analógián keresztül, amely az index hozamértéke és a diffúzió közt vonható [Plerou et al., 2000].

Egy diffúziós folyamatnál a diffundáló részecske az ütközések hatására véletlen bolyongásban vesz részt. Ha  $\Delta t$  időintervallumban az ütközések száma  $N_{\Delta t}$ , akkor a részecske által megtett út:

$$X_{\Delta t} = \sum_{i=1}^{N_{\Delta t}} \delta x_i, \quad (2.19)$$

ahol  $\delta x_i$  két ütközés közt megtett út. Egy diffundáló részecskénél az elmozdulás eloszlása,  $P(X_{\Delta t})$  Gauss lesz  $\langle X_{\Delta t}^2 \rangle = N_{\Delta t} W_{\Delta t}^2 \equiv D \Delta t$  szórással, ahol a lokális átlagos négyzetes elmozdulás,  $W_{\Delta t}^2 \equiv \langle (\delta x_i)^2 \rangle$  az individuális lépések szórása a  $\Delta t$  időintervallumban, és  $D$  a diffúziós állandó. A diffúzió esetén  $P(N_{\Delta t})$  és  $P(W_{\Delta t})$  Gauss eloszlású,  $N_{\Delta t}$  és  $W_{\Delta t}$  korrelációja rövidtávú, exponenciálisan lecsengő. Az  $\epsilon \equiv X_{\Delta t} / W_{\Delta t} \sqrt{N_{\Delta t}}$  változó korrelálatlan és normális eloszlású. Ezekből a fenti tulajdonságokból következik, hogy  $X_{\Delta t}$  korrelálatlan, Gauss eloszlású lesz.

Az index árváltozása esetében a  $\Delta t$  idő alatt történt elmozdulásnak az ezen idő alatt bekövetkezett árváltozások összege felel meg:

$$G_{\Delta t} = \sum_{i=1}^{N_{\Delta t}} \delta p_i, \quad (2.20)$$

ahol  $N_{\Delta t}$  a  $\Delta t$  idő alatt bekövetkező tranzakciók száma (az aktivitás). Hasonlóan az előbbiekhöz,  $W_{\Delta t}$  itt is az egyedi árugrások szórását jelöli. E két mennyiség eloszlását és autokorrelációját vizsgálva azt kapták, hogy: (i)  $P(N_{\Delta t} > x) \sim x^{-\beta}$ , ahol  $\beta = 3.4$ ,

(ii)  $N_{\Delta t}$  autokorrelációja hatványfüggvény lecsengésű, (iii)  $P(W_{\Delta t} > x) \sim x^{-\gamma}$ , és  $\gamma = 2.9$ , (iv)  $W_{\Delta t}$  autokorrelációja exponenciálisan cseng le. Hasonlóan a diffúziós folyamathoz, a fenti mennyiségekből képzett változó,

$$\epsilon \equiv G_{\Delta t}/W_{\Delta t}\sqrt{N_{\Delta t}} \quad (2.21)$$

normális eloszlású, rövidtávú autokorrelációval.

A fenti eredményekből az következik, hogy mivel  $G_{\Delta t}$  eloszlása hatványfüggvény,  $\alpha \simeq 3$  exponenssel, ennek oka semmiképpen nem lehet az aktivitás eloszlása, mivel  $\sqrt{N_{\Delta t}}$  eloszlásának exponense  $2\beta = 6.8$ . Tehát a lassan lecsengő hozam eloszlás eredete a  $W_{\Delta t}$  eloszlásában keresendő.

Másrésről, megfigyelhetjük, hogy a volatilitás,  $v_{\Delta t}$  a diffúziós állandóval,  $D$  áll analógiában, mivel a diffúziós állandó az  $X_{\Delta t}$  szórásával van kapcsolatban:  $D\Delta t \equiv \langle X_{\Delta t}^2 \rangle = N_{\Delta t}W_{\Delta t}^2$ , és ugyanígy a volatilitás a hozam szórásával függ össze:  $v_{\Delta t}^2 \equiv \langle G_{\Delta t}^2 \rangle = N_{\Delta t}W_{\Delta t}^2$ . A vizsgálatok azt mutatták, hogy míg  $W_{\Delta t}$  autokorrelációja gyorsan lecsengő, addig  $N_{\Delta t}$  hosszútávú autokorrelációt mutat, tehát a volatilitásnál mért hatványfüggvény lecsengésű autokorreláció oka az aktivitás autokorrelációjában keresendő.

### 2.1.3. Hozam eloszlásának skálázása

Mint azt a 2.1.1 fejezetben említettük, Mandelbrot azt a megfigyelést tette, hogy a részvények a logaritmikus hozamának eloszlása Lévy eloszlású. Mint arra korábban utaltunk, a Lévy eloszlás stabil, ami azt jelenti, hogy független, Lévy eloszlású valószínűségi változók összege is Lévy lesz. A skálázás egyenletei a (2.6) egyenletben ismertetett szimmetrikus, nulla várhatóértékű eloszlás esetén a következők alakot öltik:

$$\tilde{x} = \frac{x}{n^{1/\alpha}}, \quad (2.22)$$

$$\tilde{L}_\alpha(\tilde{x}) = L_\alpha(x) n^{1/\alpha}. \quad (2.23)$$

A logaritmikus hozam definíciója  $r_{\Delta t}(t) = \log[p(t + \Delta t)] - \log[p(t)]$  alapján a valószínűségi változók összege nem más, mint egy nagyobb időkülönbséghez tartozó hozam érték:

$$\begin{aligned} r_{\Delta t}(t) + r_{\Delta t}(t + \Delta t) &= \log[p(t + \Delta t)] - \log[p(t)] + \log[p(t + 2\Delta t)] - \log[p(t + \Delta t)] = \\ &= \log[p(t + 2\Delta t)] - \log[p(t)] = r_{2\Delta t}(t + \Delta t). \end{aligned}$$

Azaz a stabilitás azt jelentené, hogy különböző időkülönbségekhez tartozó logaritmikus hozam eloszlásfüggvények alakja megegyezne, egymásba átskálázható lenne.

### A hozam eloszlás skálázásának empirikus vizsgálata

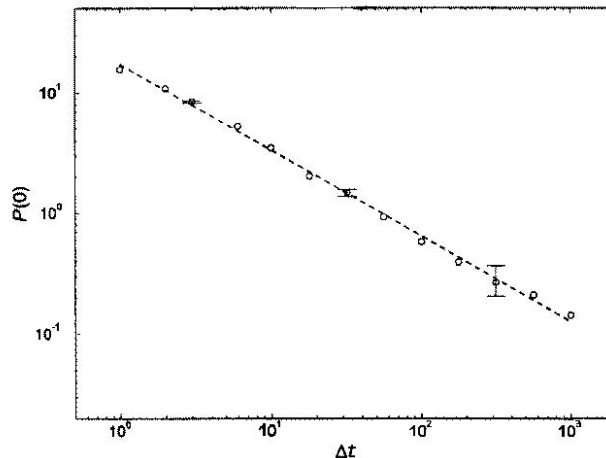
A skálázás érvényességét az empirikus adatokon Mantegna és Stanley [Mantegna and Stanley, 1995] igazolta a következő módszerrel. Mivel a szimmetrikus Lévy eloszlás sűrűségfüggvényének (a (2.6) egyenletben szereplő karakterisztikus függvény inverz Fourier transzformáltjának) nincs analitikus alakja, ezért ők a sűrűségfüggvény nulla helyen felvett értékének skálázását használták. A valószínűségrőgző függvény nulla helyen felvett értéke a következő alakú:

$$L_{\alpha, \Delta t}(x = 0) = \frac{\Gamma(1/\alpha)}{\pi\alpha(a_\alpha\Delta t)^{1/\alpha}} \quad (2.24)$$

A fenti egyenlet mindkét oldalának logaritmusát véve kapjuk, hogy

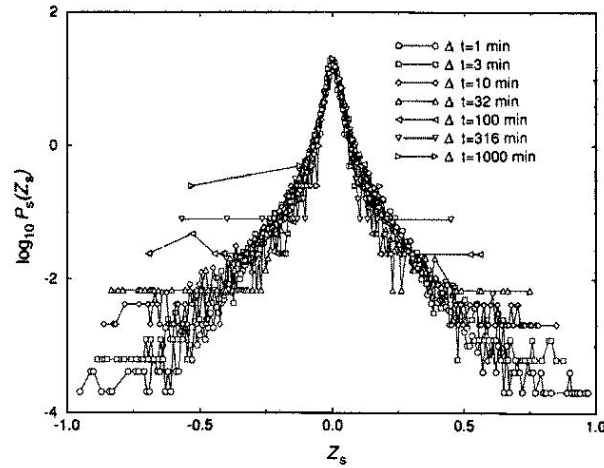
$$\ln[L_{\alpha, \Delta t}(0)] \sim -\frac{1}{\alpha} \ln[\Delta t],$$

azaz log-log skálán nézve az empirikus sűrűségfüggvény nulla helyen felvett értéke  $\Delta t$ -nek lineáris függvénye kell legyen, amit a 2.8 ábra is mutat.



2.8. ábra. A S&P500 tőzsdeindex növekményének empirikus valószínűségrőgző függvénye mint a növekményhez tartozó időkülönbség,  $\Delta t$ , függvénye log-log skálán. Az egyenes meredeksége a Lévy exponens,  $\alpha$  inverze. [Mantegna and Stanley, 1995]

A 2.8 ábrára illesztett egyenes meredekségéből az  $\alpha$  exponens értéke meghatározható,  $\alpha = 1.4 \pm 0.05$  adódott. Az  $a_\alpha$  skála paramétert a  $\Delta t = 1$  perchez tartozó empirikus sűrűségfüggvény nulla helyen felvett értékéből, illetve a (2.24) egyenletből határozták meg,  $a_\alpha = 0.00375$ . Az így kapott paraméterek behelyettesítve a (2.22), (2.23) skálaegyenletekbe a különböző időkülönbségekhez tartozó átskálázott empirikus sűrűségfüggvények valóban egy görbére esnek, ld. 2.9 ábra.



2.9. ábra. Skálázott valószínűségi sűrűség függvények lin-log skálán. A különböző skálázott sűrűségfüggvények  $1 \leq \Delta t \leq 1000$  perc időkülönbséghez tartozó növekmény sűrűségfüggvényei. A felhasznált adatok az S&P500 tőzsdeindex 1984-1989 közti nagyfrekvenciás (percnyi felbontású) értékei. [Mantegna and Stanley, 1995]

A skálázás feltétele, hogy az index-növekmény, vagy a hozam eloszlása stabil Lévy függvény legyen. Azonban mint azt a 2.1.1 fejezetben említettük, nagy hozam értékek esetén a sűrűségfüggvény gyorsabban cseng le, mint stabil Lévy eloszlás esetén. Ezért véges szórással rendelkezik, ami azt jelenti, hogy a centrális határeloszlás tétele értelmében nagy számú változó összege (azaz ebben az esetben nagy  $\Delta t$  időkülönbség) esetén a sűrűségfüggvény normális eloszláshoz kell, hogy konvergáljon. Ez a tulajdonság is teljesül [Mantegna and Stanley, 1995], de a konvergencia nagyon lassú. Az időkülönbség, amelynél a hozam eloszlása normális eloszlású lesz, néhány hónap nagyságrendbe tartozik. Felmerül a kérdés, hogy amennyiben az empirikus eloszlás véges szórású, miért lehetséges mégis a skálázása az időkülönbség több nagyságrendjén keresztül, illetve mi okozza a lassú konvergenciát.

Az egyik lehetséges oknak az tűnik, hogy a centrális határeloszlás tétele független valószínűségi változókra vonatkozik. A logaritmikusan nagy hozam értékeknél azonban láttuk, hogy ugyan korrelálatlanok, de magasabb hatványaik lassan lecsengő, hosszútávú korrelációt mutatnak, azaz nem függetlenek. Ezt az állítást, hogy a lassú konvergencia a hosszútávú magasabb rendű korrelációk következménye, igazolta [Liu et al., 1999] cikke. (i) Független valószínűségi változókat generáltak, melyek eloszlásfüggvénye szintén hatványfüggvény lecsengésű,  $P(\xi > x) \sim x^{-\alpha}$ , ugyanakkora exponenssel, mint amit a valós adatok esetén mértek,  $\alpha \simeq 3$ . Megmutatták, hogy ezek összegének eloszlása sokkal gyorsabban konvergál a normális eloszláshoz. (ii) Az S&P500 indexből alkotott hozamidősort,  $r_{\Delta t}(t)$  „összekeverték” úgy, hogy az ezáltal keletkezett új hozamidősort,  $r_{\Delta t}^{sh}(t)$  már nem rendelkezett hosszútávú korrelációval. Ezek után ezen az új idősoron vizsgálták a különböző időkülönbségekhez tartozó sűrűségfüggvényeket.

Azt tapasztalták, hogy ezek nem skálázhatóak, és sokkal gyorsabban konvergálnak a normális eloszláshoz, mint az eredeti idősor.

A lassú konvergencia másik oka esetleg abban rejlik, hogy ugyan a hozam eloszlása egy levágott Lévy eloszlás, tehát véges szórással rendelkezik, de a levágás csak nagy hozam értékeknél tapasztalható [Mantegna and Stanley, 1994]. A szerzők független, levágott Lévy eloszlású valószínűségi változókat generáltak, az alábbi eloszlásfüggvény alapján:

$$T(x) = \begin{cases} 0, & x > l, \\ cL_\alpha(x), & -l \leq x \leq l, \\ 0, & x < -l, \end{cases} \quad (2.25)$$

ahol  $l$  a levágási paraméter,  $L_\alpha(x)$  egy szimmetrikus,  $\alpha$  exponensű Lévy eloszlásfüggvény és  $c$  a normálási paraméter. Azt vizsgálták, hogy hány darab ilyen eloszlású valószínűségi változót kell összeadni,  $z_n = \sum_{i=1}^n x_i$ , hogy az összeg eloszlása Gauss legyen. Azt tapasztalták, hogy az átcsapási szám,  $n_x$ , a levágási paraméter, és az exponens alábbi függvénye:

$$n_x \sim l^\alpha$$

A fenti összefüggést analitikus levezetéssel is alátámasztották. Látszik, hogy az átcsapáshoz szükséges valószínűségi változók száma erősen nő a levágási paraméter függvényében.

#### 2.1.4. ARCH, GARCH modellek

A 2.1.2 fejezetben láttuk, hogy a logaritmikus hozam volatilitása időfüggő sztochasztikus folyamat. Egy ilyen folyamat leírására legalkalmasabb az ARCH (Autoregressive Conditional Heteroscedasticity) modell [Engle, 1982]. Ez egy olyan diszkrét idejű sztochasztikus folyamat, amelynek a feltételes szórása egy adott időpillanatban a véletlen jel négyzetének múltbeli értékétől függ. Amennyiben  $x_t$  jelöli a véletlen jelsorozatunkat, és  $\sigma_t$  ennek feltételes, időfüggő szórását, a folyamatot a következő alakban írhatjuk fel:

$$x_t \equiv \eta_t \sigma_t, \quad (2.26)$$

ahol  $\eta_t$  független, azonos eloszlású véletlen változók, nulla várhatóértékkel és egységnyi szórással.

Ezek után a  $p$  paraméterszámú ARCH( $p$ ) modell a feltételes szórásnak,  $\sigma_t$ , a múlttól való függését írja le:

$$\sigma_t^2 = \alpha_0 + \alpha_1 x_{t-1}^2 + \dots + \alpha_p x_{t-p}^2, \quad (2.27)$$

ahol  $\alpha_0, \alpha_1, \dots, \alpha_p$  pozitív paraméterek. Ahhoz, hogy az ARCH( $p$ ) folyamatot teljesen definiáljuk, meg kell határoznunk az  $x_t$  jelsorozat  $t$  időpillanatban vett feltételes valószínűségi sűrűség függvényét,  $f_t(x)$ -et, amelyet általában normális eloszlásúnál szoktak

választani. Ekkor a (2.26) egyenletből következik, hogy – mivel  $\eta_t$  és  $\sigma_t$  független változók –  $\eta_t$  normális eloszlású lesz.

Sok esetben ahhoz, hogy az ARCH(p) modell jól illeszkedjen az általunk vizsgált sztochasztikus folyamathoz, nagy  $p$  számú paraméter meghatározása lenne szükséges, ami megnehezíti a feladatunkat. Ennek elkerülésére vezette be Bollerslev [Bollerslev, 1986] az ARCH modell egy továbbfejlesztett változatát a GARCH(p,q) (generált ARCH modell) modellt:

$$\sigma_t^2 = \alpha_0 + \alpha_1 x_{t-1}^2 + \dots + \alpha_q x_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_p \sigma_{t-p}^2, \quad (2.28)$$

ahol  $\alpha_0, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p$  paraméterek, és  $x_t$  a véletlen változó, nulla várhatóértékkel, és  $\sigma_t^2$  szórással. (A gazdasági folyamatok elemzésénél általában a legegyszerűbb, a GARCH(1,1) folyamatot szokták alkalmazni.)

Könnyen belátható, a fenti definíciókat figyelembe véve, hogy az  $x_t^2$  jelsorozat időbeli autokorrelációja exponenciális lecsengést mutat

$$\langle x_t^2 x_{t+n}^2 \rangle = Ae^{-n/\tau}, \quad (2.29)$$

ahol a karakterisztikus idő,  $\tau = \frac{1}{|\ln(\alpha_1 + \beta_1)|}$ .

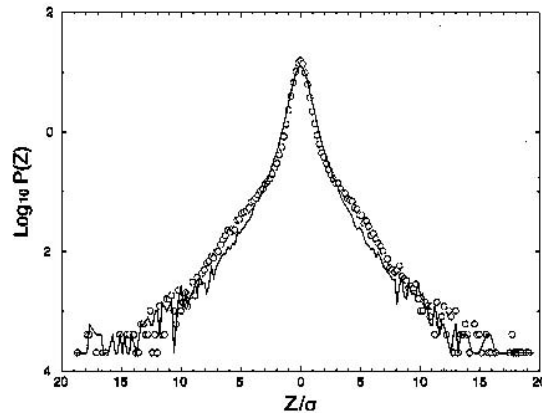
### A GARCH(1,1) modell alkalmazása az empirikus adatokra

A valós tőzsdei adatok, és egy GARCH(1,1) folyamat legszembevetőbb különbsége, hogy míg a 2.1.2 fejezetben azt láttuk, hogy a logaritmikus hozam négyzetének autokorrelációja hosszútávú, hatványfüggvény lecsengésű, addig a (2.29) egyenlet azt mutatja, hogy a GARCH(1,1) folyamatnál exponenciális lecsengésű. Ezt az ellentmondást úgy lehet a legegyszerűbben feloldani, hogy olyan  $\alpha_1, \beta_1$  paramétereket választunk, amelyek összege közel lesz egyhez. Így a karakterisztikus idő,  $\tau$ , néhány hónap nagyságrendjébe fog esni, azaz úgy tűnik, mintha hatványfüggvény lecsengésű lenne az autokorrelációs függvény [Akgiray, 1989]. (Az S&P500 adatokon végzett analízis alapján meghatározott paraméterek például  $\alpha_1 + \beta_1 = 0.98407$  értéket adtak, amelyből  $\tau = 62.3$  üzletelési nap jön ki karakterisztikus időnek [Akgiray, 1989].)

Egy másik összevetési módszer a sűrűségfüggvény és ennek skálázása a valós adatok illetve egy GARCH(1,1) folyamat esetén. [Mantegna and Stanley, 1998]. A GARCH(1,1) modell három paraméterét,  $\alpha_0, \alpha_1, \beta_1$  úgy választották meg, hogy a lehető legjobban írja le a valós adatokat, valamint a szórásnégyzete,  $\sigma^2 = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}$  és a kurtózisa,  $\kappa = 3 + \frac{6\alpha_1^2}{1 - 3\alpha_1^2 - 2\alpha_1\beta_1 - \beta_1^2}$  megegyezzen a valós adatokon mért értékekkel,  $\sigma^2 = 0.00257$  illetve  $\kappa \approx 43$ .

Azt tapasztalták, hogy az így generált GARCH(1,1) sztochasztikus folyamat feltétel nélküli sűrűségfüggvénye elég jól illeszkedik az empirikus sűrűségfüggvényre; 2.10 ábra. Ez nem azt jelenti, hogy ez a GARCH(1,1) folyamat más  $\Delta t$  értékekre is jól

jellemzi az empirikus folyamatot, hiszen ehhez tudnia kéne a skálázást is, amit a valós adatokon megfigyeltek. Hasonlóan, mint azt a 2.1.3 fejezet említi, a nulla helyen vett sűrűségfüggvény,  $P(0)$ , skálázását vizsgálták. Azt tapasztalták, hogy ugyan a GARCH(1,1) folyamat valószínűségi sűrűség függvényére is teljesül, hogy  $P(0)$  hatványfüggvénye  $\Delta t$ -nek, de a skálázási exponens értéke,  $\alpha_{GARCH} = 1.88$ , jelentősen eltér a valós adatok esetében kapott értéktől,  $\alpha \simeq 1.4$ .



2.10. ábra. Az empirikus sűrűségfüggvény és a GARCH(1,1) modell által generált folyamat feltétel nélküli sűrűségfüggvénynek összehasonlítása. Az empirikus adatok az S&P500 tőzsdeindex  $\Delta t = 1$  perces időközönbségekhez tartozó hozam értékei voltak. A GARCH(1,1) folyamat paramétereit,  $\alpha_0 = 2.3 \times 10^{-5}$ ,  $\alpha_1 = 0.09105$ ,  $\beta_1 = 0.9$  úgy választották meg, hogy a kurtózis illetve a szórásnégyzet megegyezzen az empirikus adatokon mért értékekkel. [Mantegna and Stanley, 2000]

Ha azt feltételezzük, hogy a tőzsdei árváltozás tulajdonképpen sok kis piaci befolyás egymástól független hozzájárulásának összegeként áll elő [Cont et al., 1997], természetes lenne, hogy a centrális határeloszlás tétele értelmében normális eloszlású legyen. A valós adatokon végzett vizsgálatok azonban azt mutatják, hogy a Gauss eloszlás alulbecsüli a nagy árváltozások előfordulásának valószínűségét. Felmerül a kérdés, hogy ez a leptokurtikus jelleg nem az időfüggő szórás következménye-e.

Ha a valódi adatsorra egy GARCH(1,1) folyamatot illesztünk, az illesztett paraméterek meghatározzák az időfüggő szórást (2.28), amellyel ha a (2.26) egyenlet alapján leosztjuk az eredeti idősort, tehát  $\eta = x_t/\sigma_t$ , egy normális eloszlású sztochasztikus folyamatot kell kapnunk (amennyiben az alkalmazott GARCH folyamatban a feltételes sűrűségfüggvényt Gaussnak választjuk). A vizsgálatok azonban azt mutatják, hogy noha az így keletkezett idősor sűrűségfüggvényének exponense nagyobb, mint az eredeti, de még mindig nem éri el a normális eloszlásnak megfelelő értéket,  $\alpha = 2$  [Hsieh, 1991, Bollerslev and Chou, 1992]. Ez azt jelenti, hogy egy időfüggő szórással rendelkező Gaussi sztochasztikus folyamat feltétel nélküli eloszlása hordozhat leptokurtikus jelleget, de az ARCH típusú modellek nem tudnak teljes mértékben

elszámolni a hatványfüggvény viselkedéssel.

## 2.2. Keresztkorrelációk részvényárfolyamok közt

Eddig csak egyetlen részvény, vagy a tőzsdeindex időbeli fejlődésére koncentráltunk, annak tulajdonságait vizsgáltuk. A tőzsdét azonban több részvény egyidejű változása határozza meg, amelyek egymástól nem függetlenek, egymás árfolyamát befolyásolják. Ez a fejezet azzal foglalkozik, hogy a különböző részvények hogyan hatnak egymásra.

A részvények közti kölcsönhatás vizsgálata kiemelten fontos a piaci kockázat minimalizálásának szempontjából. A minimalizálás legalkalmasabb módja egyszerre több (különböző mennyiségű) részvény vásárlása, egy *portfólió* összeállítása. Képzeljük például el, hogy létezik két részvény  $A$  és  $B$ , amelyek tökéletesen antikorreálnak, azaz ha az  $A$  részvény árfolyama emelkedik, a  $B$  részvényé ugyanabban az időpillanatban ugyanannyival süllyed, és fordítva. Ebben az esetben a vásárlónak, aki mindkét részvényt megvette, nulla kockázata lesz. Ez természetesen egy idealizált eset, amely a valós piacon sohasem alakul ki, de általánosan igaz, hogy a diverzifikáció által csökkenthetjük a kockázatot. A fenti példában azonban nem csak a kockázat de a várható haszon is nulla lesz. A helyes kérdés feltevése tehát az, hogy egy adott hozamszint (nyereség) mellett mekkora a kockázat. Az optimális portfólió számításának módszerére itt nem térünk ki, ez megtalálható [Bouchaud and Potters, 1997, Elton and Gruber, 1995] könyvekben.

A részvények közti kölcsönhatás vizsgálatára a legegyszerűbb módszer a különböző részvények árfolyamváltozása közt mért keresztkorreláció analízise, hiszen ez egy közvetlenül mérhető mennyiség:

$$\rho_{ij} = \frac{\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle}{\sqrt{\langle (r_i - \langle r_i \rangle)^2 \rangle \langle (r_j - \langle r_j \rangle)^2 \rangle}}, \quad (2.30)$$

ahol  $\langle \cdot \rangle$  időbeli átlagolást jelent a részvény árfolyamok teljes vizsgált időtartományára.

### A korrelációk és a kovariancia mátrix tulajdonságai

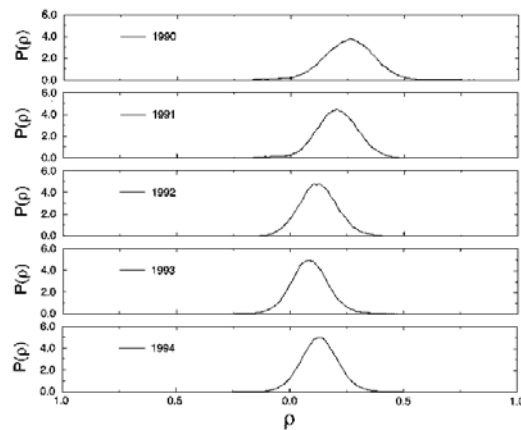
A (2.30) egyenlet által definiált, különböző részvények árfolyamváltozásai közt mért egyidejű korreláció tulajdonságait vizsgálták két nagyobb részvényhalmazon, öt egymást követő évben [Mantegna and Stanley, 2000]. (Az egyidejű szó arra utal, hogy a  $\langle r_i r_j \rangle$  szorzatban ugyanahhoz a pillanathoz tartozó hozam értéket vizsgáljuk a két részvény,  $i, j$  esetén.) Az első részvényhalmaz a Dow Jones Industrial Average (DJIA) volt, amelybe 30 vállalat tartozik, a második a már korábban is említett Standard and Poor's 500 (S&P500) 500 vállalattal. Mindkét esetben azt tapasztalták, hogy a keresztkorrelációs együtthatók közül legnagyobb pozitív érték messze meghaladja a

legnagyobb negatív értéket, azaz az antikorrrelációs effektusok sokkal kisebbek. Ezt mutatja a 2.1 táblázat.

2.1. táblázat. Az S&P 500 vállalatai közt, 1 éves időintervallumra mért keresztkorrrelációs együtthatók,  $\rho_{i,j}$  minimum és a maximum értékei a különböző években. [Mantegna and Stanley, 2000]

Idő	Min.	Max.
1990	-0.30	0.81
1991	-0.29	0.74
1992	-0.25	0.73
1993	-0.27	0.81
1994	-0.25	0.82

Az S&P500 esetén a keresztkorrrelációs együtthatók száma ( $499 \cdot 500 / 2$ ) már megfelelő statisztikát nyújt az együtthatók eloszlásának,  $P(\rho_{ij})$  vizsgálatához. Az eredmény, 2.11 ábra, egyrészt igazolja, hogy a negatív korrrelációk súlya sokkal kisebb, mint a pozitívoké, másrészt az is leolvasható, hogy míg a várhatóérték erősen változik az idő függvényében, addig a szórás nagyjából konstansnak vehető.



2.11. ábra. Korrelációs együtthatók eloszlása 5 egymást követő évben. Az együtthatókat az S&P500 index részvényei közt mérték, egy év időintervallumra átlagolva. [Mantegna and Stanley, 2000]

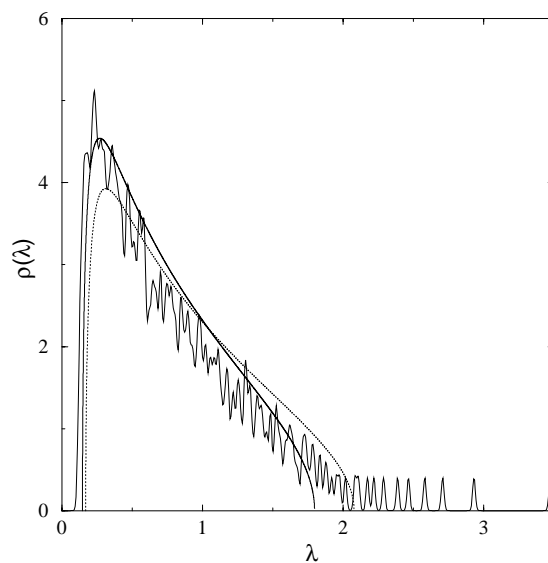
A részvények közti kölcsönhatást a logaritmikus hozamaik közt mért korrrelációval írjuk le, mivel ez az egyik legegyszerűbben mérhető mennyiség. Egy  $N$  elemű részvényhalmaz esetén a kovariancia mátrixot (a  $\rho_{ij}$  együtthatókból felépített mátrix)  $N$  darab, egyenként  $T$  hosszúságú idősorból kell meghatározni. Ha  $T$  nem sokkal nagyobb mint  $N$ , akkor azt várjuk, hogy a kovariancia mátrix meglehetősen zajos lesz, ezért az alkalmazásakor körültekintően kell eljárni. Lényeges, hogy meg tudjuk

különböztetni azon sajátértékeket illetve sajátvektorokat, amelyek valós információt hordoznak, azoktól, amelyeket csak a zaj okozza, és időben instabilak. Ennek érdekében azt vizsgálták [Laloux et al., 1998, Plerou et al., 1999], hogy a kovariancia mátrix mennyire különbözik egy véletlen mátrixtól, amelynek elemei korrelálatlan idősorok közt mért keresztkorrelációs együtthatók. Azt találták, hogy az empirikus korrelációs mátrix sajátértékeinek sűrűsége,  $\rho(\lambda) = \frac{1}{N} \frac{dn(\lambda)}{d\lambda}$ , (ahol  $n(\lambda)$  a  $\lambda$ -nál kisebb sajátértékek száma), jól leírható a véletlen mátrix elméletből [Sengupta and Mitra, 1997] kapott sűrűséggel, 2.12 ábra:

$$\rho(\lambda) = \frac{Q}{2\pi\sigma^2} \frac{\sqrt{(\lambda_{max} - \lambda)(\lambda - \lambda_{min})}}{\lambda}, \quad (2.31)$$

$$\lambda_{min}^{max} = \sigma^2(1 + 1/Q \pm 2\sqrt{1/Q}),$$

ahol  $Q = T/N$  az idősorok hosszának, és a vizsgált részvények számának hányadosa, és  $\sigma^2$  az idősorok szórásának átlaga. Az empirikus kovariancia mátrix sajátértékeknek csak 6%-a nagyobb, mint  $\lambda_{max}$ , ezek a releváns, információt hordozó sajátértékek.



2.12. ábra. Az empirikus kovarianciamátrix sajátérték sűrűségfüggvénye. A kovariancia mátrix az S&P500 tőzsdeindex  $N = 406$  részvényeinek keresztkorrelációját tartalmazza, melyek  $T = 6$  évre vannak átlagolva. (Ezekből  $Q = T/N = 3.22$ ). A szaggatott vonal az (2.31) egyenlet alapján számított függvény, a  $Q = T/N = 3.22$  és  $\sigma^2 = 0.85$  paraméterek felhasználásával. A sima vonal egy ennél jobb fitt eredménye, amelyet egy kicsit kisebb szórásnégyzet,  $\sigma^2 = 0.74$  alkalmazásával kaptak [Laloux et al., 1998].

Azt, hogy a kis sajátértékek a zajnak felelnek meg, úgy a legkönnyebb belátni, ha megvizsgáljuk a hozzájuk tartozó sajátvektorok struktúráját. Amennyiben a  $\lambda_\alpha$  sajátértékhez tartozó  $v_\alpha$  sajátvektor nem hordoz információt, úgy a komponenseiből

képzett eloszlás az úgynevezett Porter-Thomas eloszlásnak felel meg [Mehta, 1995]:

$$P(u) = \sqrt{\frac{1}{2\pi}} \exp\left\{-\frac{u^2}{2}\right\}$$

alakú, ahol  $u = v_{\alpha,i}$  a sajátvektor komponenseit jelöli. Az analízis azt mutatta, hogy  $\lambda_{max}$ -nál kisebb sajátértékekhez tartozó sajátvektorokra a fenti eloszlás jól illeszkedik, ugyanakkor az ennél nagyobb sajátértékekhez tartozó sajátvektorok komponenseinek eloszlása szignifikáns eltérést mutat.

### 2.2.1. Csoportosítási módszerek

Részvények árfolyamváltozásai közt mért keresztkorreláció mérése egyrésről fontos szerepet játszik a portfólió elméletben az optimális portfólió kiszámításához, másrésről megvizsgálhatjuk azt is, hogy amennyiben a korrelációkat mint valami effektív erőket feltételezzük a részvények közt, hogyan hatnak ezek egymásra, azaz hogyan lehet ez alapján csoportosítani a tőzsdei részvényeket. A következő alfejezetekben két általános csoportosítási módszert mutatunk be, amelyekkel a részvények közti hierarchikus rendezés elvégezhető.

#### Szuper-paramágneses klasztereződés

Az adatok csoportokba rendezése egy általános probléma. Vegyünk egy  $N$  elemű halmazt. Úgy szeretnénk az elemeket csoportokba rendezni, hogy a csoportok az adatokban rejlő természetes osztályokat adják vissza. (Vegyünk egy egyszerű példát. Tegyük fel, hogy a halmazunk különböző állatokot tartalmaz. Ekkor elvárjuk, hogy a végső csoportosításnál az azonos fajtához tartozó állatok kerüljenek egy csoportba.) Azaz úgy kell létrehozni a csoportokat, hogy a csoporton belüli elemek jobban hasonlítsanak egymásra, mint a más csoportokban lévő elemekre. A hasonlóság definiálására a legegyszerűbb példa, ha minden elemnek megmérjük  $D$  darab tulajdonságát, így minden elemnek megfelel egy pont egy  $D$  dimenziós térben. A hasonlóság a pontok közti távolsággal egyezik meg.

Kétféle csoportosítási eljárás létezik. Az első a *paraméteres* csoportosítás, ahol eleve feltételezünk valamilyen struktúrát a csoportokra vonatkozóan, azaz létezik egy globális kritérium, és a pontok csoportokba osztása ennek minimalizálásával történik. Ha azonban nem létezik ilyen *a priori* ismeretünk, akkor a másik, a *paraméter nélküli* csoportosítást kell használnunk, amely lokális kritérium alapján rendezi az elemeket.

A paraméter nélküli csoportosítási módszerhez tartozik a *szuper-paramágneses klasztereződés* [Blatt et al., 1996], amely egy inhomogén,  $q$  állapotú Potts modell egyensúlyi állapotában az egy irányba mutató spineket sorolja egy csoportba. A spin-párok vizsgálata egy lokális kritériumnak felel meg. A csoportosítási feladatnak a statisztikus fizikai megközelítése elég természetesen adódik.

Minden eljárásnál a csoportok azonosítása általában valamilyen célfüggvény minimalizálásának az eredménye. Lehetséges azonban, hogy ugyanahhoz a célfüggvény értékhez több konfiguráció is tartozik. Ekkor mi alapján tüntetnénk ki valamelyiket? A Potts modell esetén nem tüntetjük ki semelyiket, hanem mindegyiket egyforma súllyal választjuk a mikro-kanonikus sokaságnak megfelelően.

A másik probléma a felbontás kérdése. Egy halmazt nem csak egyféleképpen lehet csoportosítani. Visszatérve az állatos példára, az összes elemet tekinthetjük egy csoportnak, (mindegyik elem az állatok összegző kategóriájába tartozik), tekinthetjük a fajokat, de lehetséges minden egyedat külön csoportba sorolni. A Potts modellben a hőmérséklet játssza a felbontás szerepét (kanonikus sokaságban). Nulla hőmérsékleten minden spin egy irányba fog mutatni – *ferromágneses fázis* –, ez megfelel annak, ha minden elemet egy csoportba sorolunk. Nagyon nagy (végtelen) hőmérsékleten a fluktuációk lerombolnak minden rendet, a spinek függetlenné válnak – *paramágneses fázis* –. Ez annak a jelenségnek felel meg, amikor minden elem saját csoportot alkot. Ha megfelelően választjuk a modell paramétereit, e két extrém fázis közt létezik egy harmadik állapot, a *szuper-paramágneses fázis*, amelyben léteznek spin-csoportok, melyek egy irányba mutatnak, de a különböző csoportok egymástól függetlenek. Ilyen módon a hőmérséklet függvényében egy hierarchikus rendezést kapunk. Ahogy növeljük a hőmérsékletet, úgy esnek szét a csoportok egyre kisebb halmazokra.

A modell ezek után a következő: a halmaz minden elemének feleljen meg egy spin amely  $q$  különböző értéket vehet fel,  $\{s_i\}_{i=1}^N = 1, \dots, q$ . A Potts modell Hamilton függvénye:

$$\mathcal{H}[\{s\}] = - \sum_{(i,j)} J_{ij} \delta_{s_i, s_j}, \quad (2.32)$$

ahol  $\delta_{s_i, s_j}$  a Kronecker szimbólum és  $J_{ij}$  az  $i$ -dik és a  $j$ -dik spin közti csatolási állandó, amely – ha a hasonlóságot, mint a  $D$  dimenziós „tulajdonság térben” vett távolságot definiáljuk – monoton csökkenő függvénye a távolságnak.

Mivel minden elem valamilyen szinten kölcsönhat minden másikkal, azaz a (2.32) egyenletben az összegzés alatt szereplő  $(i, j)$  jelölés az összes párra való összegzést jelent. A csatolási állandónak a távolságtól való függése többféleképpen definiálható. A szerzők [Blatt et al., 1996] az alábbi javaslatot tették:

$$J_{ij} = \frac{1}{K} \exp \left\{ -\frac{d_{ij}^2}{2a^2} \right\}, \quad (2.33)$$

ahol  $d_{ij}$  az  $i$  és  $j$  pontok közti távolság,  $a = \frac{1}{N} \sum_i \min_j (d_{ij})$  az átlagos legkisebb távolság a halmaz elemei közt, és  $K$  az  $a$  távolságon belüli elemek száma. Ezáltal a csatolási állandót úgy definiálták, mintha elsőszomszéd kölcsönhatás lenne az elemek közt, azaz a hosszú hatótávú kölcsönhatást rövid hatótávolságúvá tette. Ez azért is szükséges, mert ha a csatolási állandó nem vág le elég gyorsan, nem alakul ki a szuper-paramágneses fázis.

A csoportok azonosításához az egyensúlyi spin-spin korrelációs függvényt,  $\langle \delta_{s_i, s_j} \rangle$  kell meghatározni (ahol  $\langle \cdot \rangle$  a termodinamikai átlagot jelöli). Ha a korreláció meghalad egy küszöböt:  $\langle \delta_{s_i, s_j} \rangle > \theta$ , akkor az  $i$  és a  $j$  elemek egy csoportba tartoznak. A spin-spin korrelációs függvény eloszlását vizsgálva azt láthatjuk, hogy két csúcsa van, az egyik az  $1/q$ , a másik az 1 értéknél. Ezek között nullához közeli értéket vesz fel, ezért az eredmények nem túl érzékenyek a  $\theta$  érték megválasztására. A szerzők a  $\theta = 0.5$  küszöbértéket javasolták [Blatt et al., 1996].

Az egyensúlyi spin-spin korrelációs függvény meghatározása Monte-Carlo szimulációval történik. A legalkalmasabb ehhez a Swendsen-Wang klaszter algoritmus [Swendsen and Wang, 1987] alkalmazása, mert ennél a módszernél a fázisátalakulás pontjában nem lép fel a kritikus lelassulás jelensége. Mielőtt azonban meghatározzuk a spin-spin korrelációs függvényt, meg kell vizsgálni, hogy milyen hőmérsékleten szimuláljuk a rendszert, azaz melyek lesznek azon hőmérsékletek, ahol a szuperparamágneses fázisok megjelennek. A fázisátalakulás pontjában a szuszeptibilitás,  $\chi(T)$  értéke megnő, a különböző fázisokat tehát ezen csúcsok választják el egymástól. A szuszeptibilitást a rendparaméter,  $m$ , fluktuációja adja meg:

$$\chi = \frac{N}{T} (\langle m^2 \rangle - \langle m \rangle^2). \quad (2.34)$$

A rendparaméter többféleképpen definiálható. A szerzők [Blatt et al., 1996] az alábbi javasolták:

$$m = \frac{(N_{max}/N)q - 1}{q - 1}, \quad (2.35)$$

ahol  $N_{max} = \max\{N_1, N_2, \dots, N_q\}$  és  $N_i$  az  $i$  irányba álló spinek száma.

Ha mérjük tehát a szuszeptibilitást, mint a hőmérséklet függvényét,  $\chi(T)$ -t, meghatározhatjuk a hőmérséklet tartományokat, ahol a szuperparamágneses fázis előfordul. Ezeken a hőmérsékleten szimulálva a rendszert, az egyensúly elérése után, a spin-spin korrelációs függvény mérésével lehet meghatározni, hogy mely elemek kerülnek egy csoportba.

### Minimális kifeszítő fa

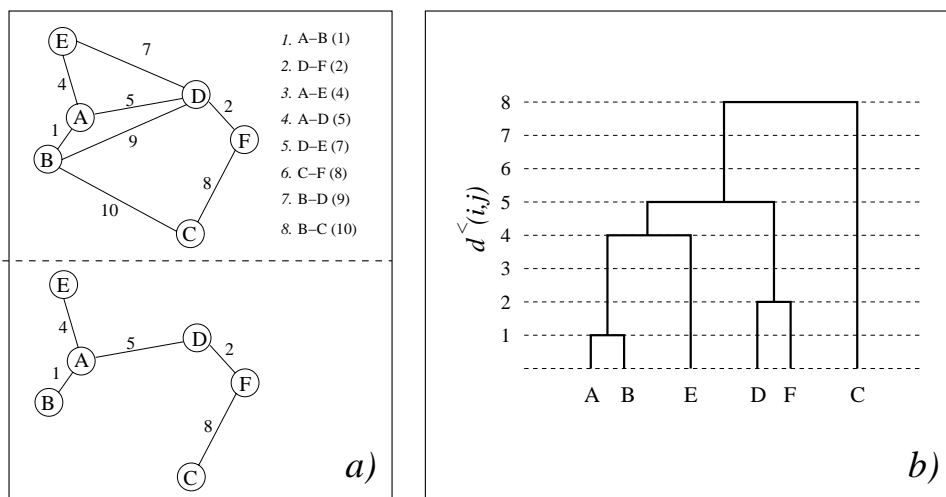
Tekintsünk egy  $G$  gráfot,  $V$  darab csomóponttal és  $E$  darab súlyozott éllel. (A súlyozás pl. lehet az él Euklideszi távolsága.) Egy *kifeszítő fa*  $G$ -nek olyan alhálózata, amely minden csomóponton legalább egyszer átmegy. A *minimális kifeszítő fa* (MST) ezek közül az, amelyre az él súlyainak (hosszainak) összege minimális.

Ahhoz, hogy a részvények esetében megszerkesszük a MST-t először is definiálnunk kell két részvény közti távolságot,  $d(i, j)$ , [Mantegna, 1999]:

$$d(i, j) = \sqrt{2(1 - \rho_{ij})}, \quad (2.36)$$

ahol  $\rho_{ij}$ , a (2.30) egyenlet által definiált keresztkorreláció. Az egyenletben szereplő  $r_i$  értékek az  $i$ -dik részvény egy nap időkülönbséghez tartozó logaritmusos hozam értéke. Könnyen belátható, hogy a (2.36)-ben definiált mennyiség kielégíti a metrikus távolság axiómáit: (i)  $d(i, j) = 0$ , ha  $i = j$ ; (ii)  $d(i, j) = d(j, i)$ ; (iii)  $d(i, j) \leq d(i, k) + d(k, j)$ . Mivel  $\rho_{ij} = 1$ , ha a két idősor teljesen megegyezik, ebből következik (i). A korrelációk szimmetriájából  $\rho_{ij} = \rho_{ji}$  következik, hogy (ii) teljesül. Végül ha az  $r_i$  hozam idősort úgy tekintjük, mint egy vektort, amelynek minden eleme az idősor egy tagja, akkor (2.36) nem más mint az ezekből képzett normált vektorok,  $\tilde{\mathbf{r}}_i = \frac{\mathbf{r}_i - \langle \mathbf{r}_i \rangle}{\sqrt{\langle (\mathbf{r}_i - \langle \mathbf{r}_i \rangle)^2 \rangle}}$ , Euklideszi távolsága:

$$d^2(\tilde{\mathbf{r}}_i, \tilde{\mathbf{r}}_j) = (\tilde{r}_{i,1} - \tilde{r}_{j,1})^2 + \dots + (\tilde{r}_{i,T} - \tilde{r}_{j,T})^2 = \tilde{\mathbf{r}}_i^2 + \tilde{\mathbf{r}}_j^2 - 2 \sum_{k=1}^T \tilde{r}_{i,k} \tilde{r}_{j,k} = 2 - 2\rho_{ij}.$$



2.13. ábra. Példa a MST generálására. Az **a)** ábrán felül elhelyezkedő gráfnak akarjuk meghatározni a MST-jét. Az éleken feltüntetett számok az élek súlyai, ezeknek a növekvő rendbe tett sorozata látható a gráf mellett, amely egyben a MST éleinek behúzási sorrendje is. Láthatjuk, hogy a második lépésben (D-F vonal) két algráfunk keletkezik – a (ii) lépés –, amelyeket csak a 4. lépésben (A-D vonal) kötünk össze. Másrésztől láthatjuk, hogy az 5., a 7. és a 8. lépésben (D-E, B-D és B-C vonal) nem húztunk be új vonalt, mert azon csomópontok, amelyeket ezek összekötnének, már össze vannak linkelve korábbról. A **b)** ábrán a hierarchikus fa látható, amelyet a csomópontok közti ultrametrikus távolságokból kapunk.

A távolság mátrix, **D**, meghatározása után a MST felépítése nagyon egyszerű; ld. 2.13 ábra. A távolság mátrix elemeit növekvő sorrendbe rendezzük. A legkisebb távolsághoz tartozó párt összekötjük, majd sorba vesszük az egyre nagyobb távolságokhoz tartozó párokat. (i) Ha olyan pár következik, amelynek az egyik tagja már szerepelt, akkor ahhoz az elemhez kötjük a pár másik tagját. (ii) Ha egy olyan pár következik, amely közül még semelyik nem szerepelt a MST-ben, akkor az egy új

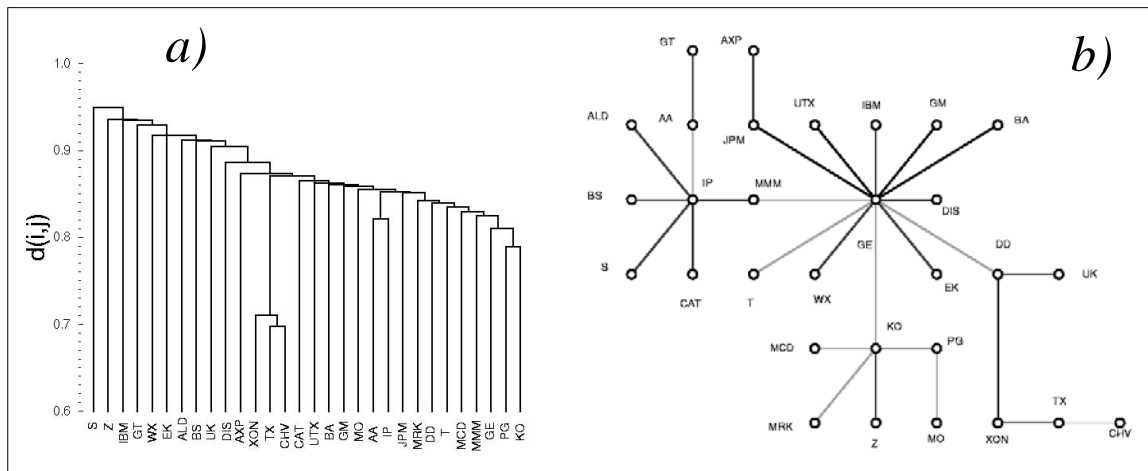
csoportot fog alkotni. (iii) Amennyiben olyan pár következik, amelynek már mindkét tagja szerepelt, és mindkét tag egy csoporton belül volt, akkor azt a párt már nem kell összekötni, mint például a 2.13/a) ábrán az 5. lépésnél a  $D - E$  szakasz behúzásánál.

A MST megszerkesztésével megalkothatjuk a távolság mátrixunkból a *szubdomináns ultrametrikus távolság mátrixot*,  $\mathbf{D}^<$ , amelynek elemeire,  $d^<(i, j)$ , fennáll, hogy a maximuma az  $i$ -ből  $j$ -be vezető élek súlyainak. Példa erre a 2.13/a) ábrán  $d^<(E, D) = \max(4, 5) = 5$ , azaz a szubdomináns ultrametrikus térben  $D$  ugyanolyan távolságra van  $E$ -től, mint  $A$ -tól, ami az eredeti gráfon nem teljesül. Az Euklideszi és az ultrametrikus tér közti fő különbség a háromszög szabályban van:

$$\begin{aligned} d(A, C) &\leq d(A, B) + d(B, C) \\ d^<(A, C) &\leq \max[d^<(A, B), d^<(B, C)] \end{aligned} \quad (2.37)$$

Az ultrametrikus tér háromszög szabálya tehát egy erősebb feltételt ad. Ez egy olyan tér, ahol csak egyenlő oldalú és egyenlő szárú háromszögek vannak, de nem létezik háromszög, amelynek mindhárom oldala különbözik.

Az ultrametrikus mátrix felhasználásával felépíthetjük a *hierarchikus fát*, 2.13/b) ábra, amely azt mutatja meg, hogy a különböző csoportok hogyan rendeződnek hierarchiába [Rammal et al., 1986].



2.14. ábra. A DJIA részvényhalmaz **a)** minimális kifizető fája, illetve **b)** hierarchikus fája [Mantegna, 1999].

Ezt a módszert használta Mantegna [Mantegna, 1999] az S&P500 illetve a DJIA részvény halmazok hierarchikus csoportokba rendezéséhez. A DJIA esetén kapott MST és a hierarchikus fa elrendezést a 2.14 ábra illusztrálja. (Az S&P500 olyan sok részvényt tartalmaz, hogy annak illusztrálása nehézségekbe ütközik.) Az ábráról az olvasható le, hogy 3 nagyobb csoport azonosítható. XON - TX - CHV (Exxon - Texaco - Chevron); PG - KO (Procter and Gamble - Coca-Cola); AA - IP (Alcoa -

International Paper). A többi vállalat ezekhez a csoportokhoz kapcsolódik. Általánosan igaz, (ezt mutatták az S&P500-on végzett vizsgálatok is) hogy azon vállalatok kapcsolódnak egy csoportba, amelyek a gazdaságban egy ágazatban vannak.

Ebben a témakörben sok tanulmány született. A Mantegna féle eredményeket csak azért emeltük ki, mert az volt az első ilyen jellegű vizsgálat, de vizsgálták ugyan ezeket az idősorokat más módszerekkel is [Giada and Marsili, 2001], illetve a fent említett MST módszert alkalmazták nem részvény, hanem különböző tőzsdeindexek idősorai közt mért keresztkorrelációkra [Bonanno et al., 2000], mivel nyilvánvalóan a különböző tőzsdék sem függetlenek egymástól.

## 2.3. Tőzsdei sok ügynökös modellek

A tőzsdei folyamatok leírására szolgáló modellek egyrésztől azon ismert sztochasztikus folyamatok – levágott Lévy folyamat, ARCH-GARCH heteroszkedasztikus folyamat –, amelyek rendelkeznek a tőzsdei árfluktuáció bizonyos karakterisztikus tulajdonságaival. Másrésztől leírhatjuk a tőzsdét úgynevezett „*mikroszkopikus*” modellekkel is, amely alatt azt értjük, hogy magát a tőzsdén lezajló árki alakítási folyamatot kíséreljük meg szimulálni úgy, hogy az ezáltal létrehozott árfluktuáció megegyezzen a valós árváltozások tulajdonságaival. Nyilvánvalóan az ár kialakításának mechanizmusa nagyon bonyolult folyamat, még ha ismert is lenne minden lépése, a szimulálásánál lehetetlen lenne az összes tényezőt figyelembe venni. Ezért ki kell emelni azon jellegzetes momentumokat, amelyek úgy tűnnek, hogy jelentősen befolyásolják a folyamatot.

Az utóbbi években sok ilyen típusú mikroszkopikus modell született, [Lux and Marchesi, 1999, Cont and Bouchaud, 2000, Maslov, 2000] más-más jellegzetességeit ragadva meg az ár kialakítási folyamatnak. Ezek általában *egy-részvényes, sok-ügynökös* modellek, azaz csak egyetlen részvény árának kialakulását határozza meg az  $N$  darab tőzsdei vásárló.

### 2.3.1. „Csorda”-hatás

Az egyik talán legegyszerűbb modell a Cont és Bouchaud által kidolgozott úgynevezett „*csordaszellemű*” piaci modell [Cont and Bouchaud, 2000]. A modell alap gondolata, hogy a piaci résztvevők nem önállóan üzemelnek, hanem csoportokba szerveződnek, és egy csoporton belül mindenki ugyanazt a üzletelési stratégiát követi, azaz olyan, mint egy birkanyáj, ahol az össze birka arra megy, amerre a vezérkos irányítja a nyáját. A feltevés az, hogy ez a csordaszellem okozza, hogy a hozam eloszlásában az extrém események olyan nagy valószínűséget kapnak.

A modell a következőképpen fogalmazható meg. Adott egy részvény, amelynek ára  $t$  időpillanatban  $p(t)$ . Adott továbbá  $N$  résztvevő, akik csoportokba rendeződnek, de a csoportok mérete,  $s$ , eltér egymástól. Egy csoporton belül mindenki ugyanazt

a  $\phi \in \{-1, 0, 1\}$  stratégiát követi, azaz elad ( $\phi = -1$ ), nem üzletel ( $\phi = 0$ ), vagy vásárol ( $\phi = 1$ ). A különböző csoportok azonban egymástól függetlenül döntenek. Amennyiben egy csoport aktív, tehát vásárol vagy elad, akkor minden tagja egyetlen részvényt vesz vagy ad el. A modell feltételezi, hogy az ár változása arányos a nettó üzletelési igénnyel, azaz a vásárlások és az eladások teljes különbségével. Ekkor a logaritmikus hozam  $r(t) = \ln[p(t)/p(t-1)]$  a következőképpen írható fel:

$$r(t) \sim \sum_{i=1}^n s_i \phi_i(t), \quad (2.38)$$

ahol  $n$  a csoportok száma,  $\phi_i(t)$  az  $i$ -edik csoport stratégiája a  $t$  időpillanatban, és  $s_i$  az  $i$ -edik csoport mérete. (Feltételezzük, hogy a csoportok mérete nem változik az idő folyamán.) Egy csoport stratégiáját minden időlépésben véletlenszerűen határozzuk meg. Egy csoport  $a$  valószínűséggel vásárol, és ugyanakkora  $a$  valószínűséggel ad el, illetve  $(1-2a)$  valószínűséggel nem üzletel.

A csoportok kialakulása hasonló egy véletlen hálózathoz. Adva van  $N$  pont, ami megfelel az  $N$  résztvevőknek. Két pont  $i$  és  $j$  közt  $p_{ij}$  valószínűséggel létesül kapcsolat. Ezeket összekötjük. Azon pontok tartoznak egy csoportba, amelyek össze vannak kötve. Az egyszerűsítés érdekében a kapcsolat kialakításának valószínűsége legyen független  $i$  és  $j$ -től, azaz  $p = p_{ij}$ . Mivel egy pont így átlagosan  $(N-1)p$  ponttal kerül kapcsolatba, és mivel minket az  $N \rightarrow \infty$  határeset érdekel, válasszuk a valószínűséget úgy, hogy véges csoportméretek alakuljanak ki, azaz legyen  $p = c/N$ . A csoportok eloszlását tehát egyetlen paraméterrel,  $c$ -vel jellemezhetjük. A véletlen hálózatok elmélete alapján [Erdős and Rényi, 1960] megmutatható, hogy  $c = 1$  esetében a csoport méreteloszlásának sűrűségfüggvénye hatványfüggvény lecsengést mutat:

$$P(s) \underset{s \rightarrow \infty}{\sim} \frac{A}{s^{5/2}} \quad (2.39)$$

Ha egy adott időpillanatban az aktív (üzletelő) csoportok száma kicsi, akkor a hozam eloszlása,  $P(r)$  hasonló viselkedést mutat a csoport méreteloszláshoz,  $P(s)$ . (Például ha feltételezzük, hogy csak egyetlen aktív csoport létezik minden időpontban, akkor a (2.38) egyenletben az összeg helyett csak egyetlen  $s_i$  tag áll, azaz ebben az esetben a két eloszlás teljesen megegyezik.) Mivel  $P(s)$  hatványfüggvény lecsengésű, kis időkülönbség esetén tehát a hozam eloszlása is hatványfüggvény lecsengésű lesz, ezzel igazolva a feltevést, hogy a csordaszellem lehet az oka az extrém események nagy előfordulási valószínűségének.

A modell azonban nem ad jó exponens értéket, hiszen (2.39) alapján a valószínűség-sűrűség függvényben az exponens értéke  $\alpha = 2.5$ , ami nem egyezik az empirikus vizsgálatok által adott  $\alpha \simeq 4$  értékkel [Gopikrishnan et al., 1999]. Másrésztől a modell nem tud számot adni a hozam abszolút értékénél megfigyelt hosszútávú autokorrelációról [Gopikrishnan et al., 1999].

### Perkolációs modell

Ezen problémák megoldására született a fenti modell egy továbbfejlesztett változata [Stauffer and Penna, 1998], amely a csoportokat mint perkolációs fürtöket definiálja. Egy rács pontjait  $p$  valószínűséggel betöltjük, a csoportokat az egymással szomszédos rácsponthalmazok fogják alkotni. A csoport méreteloszlása – feltéve, hogy a perkolációs kritikus pontban vagyunk,  $p = p_c$  – hatványfüggvény lecsengésű lesz:

$$n_s \simeq s^{-\tau}, \quad (2.40)$$

ahol az exponens értéke a perkolációs rács dimenziójától függően  $2 < \tau < 2.5$ , ha a dimenzió  $1 < d < \infty$  értékek közt változik.

A csoportok kialakítását leszámítva az árkialakítási folyamat megegyezik a Cont-Bouchaud féle modellben leírtakkal: minden csoport  $2a$  valószínűséggel lesz aktív ( $a$  valószínűséggel vásárol, és ugyanakkora valószínűséggel ad el). A különbség a Cont-Bouchaud modellhez képest, hogy itt a csoportok méretei változnak a folyamat során, méghozzá oly módon, hogy minden időlépés után az üzletelők (betöltött rácsponthalmazok) 1%-a véletlenszerűen elmozdul egy szomszédos, betöltetlen rácsponthalmazba. Azaz a csoportméretek időben korreláltak lesznek, a csoportméret autokorrelációja – és így a hozam abszolút értékének autokorrelációja is – exponenciálisnál lassabb lecsengést fog mutatni [Stauffer and Penna, 1998].

Ahhoz, hogy a hozam eloszlásának exponense megegyezzen az empirikus értékkel, további módosításokra van szükség, amik [Stauffer and Sornette, 1999] javaslatára a következők.

Elvetették azt a feltevést, hogy a hozam arányos legyen a nettó követelésekkel,  $\Delta$  (összes vásárlás és eladás összege), ehelyett egy nemlineáris összefüggést használtak:

$$r(t) \sim \sqrt{\Delta}. \quad (2.41)$$

Ez az összefüggés abból az elképzelésből származik, hogy egy  $s$  méretű üzlet lebonyolításához az  $s$ -sel arányos időre van szükség és ez alatt az idő alatt diffúziós folyamatnak megfelelő árfluktuációk lépnek fel. Ezek alapján határozta meg Zhang a (2.41) egyenletben szereplő gyökös kifejezést [Zhang, 1999].

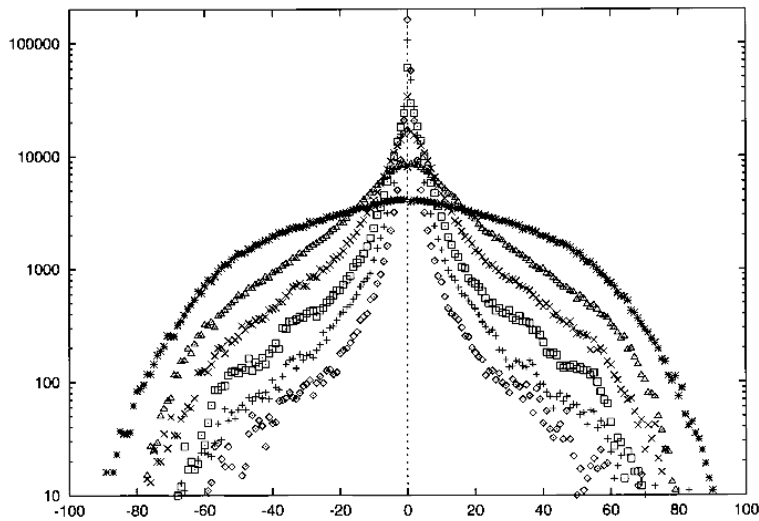
Másrésztől azt láttuk, hogy a perkolációs modell csak abban az esetben ad hatványfüggvény lecsengésű méreteloszlást, ha a csoportok kialakításánál a kritikus pontban  $p = p_c$  van a perkolációs rendszer. Azonban nincs semmi okunk, hogy ezzel a feltételezéssel éljünk. Létezik azonban egy alternatív mechanizmus is, amely akkor is hatványfüggvény lecsengést eredményez, ha nem vagyunk a kritikus pontban. Feltételeztük, hogy a rács betöltési valószínűsége,  $p$  a folyamat során nem változik. A  $p$  valószínűség a csoport kialakításának erősségét jellemezi. A valóságban azonban könnyen előfordulhat, hogy időben változik a csordaszellem erőssége. Amennyiben tehát  $p$  időbeli állandóságának feltételét elvetjük, és értékét minden időlépésben véletlenszerűen választjuk meg, akkor a hozam eloszlását a különböző  $p$  értékekre kapott

eloszlások  $p$ -re való átlagolásával kapjuk meg:  $P(r) = \langle P(r, p) \rangle_p$ . (Elméleti megfontolások alapján a véletlenszerű betöltés értékét csak  $0 < p < p_c$  közt választjuk.)

Ha mindkét módosítást – a nemlineáris hozamfüggvény illetve időfüggő betöltési valószínűség – figyelembe vesszük, akkor az így meghatározott exponensre,  $\mu \simeq 3.9$  értéket kapunk, ami jó egyezést mutat az elméleti megfigyelésekkel.

### 2.3.2. Aktivitás

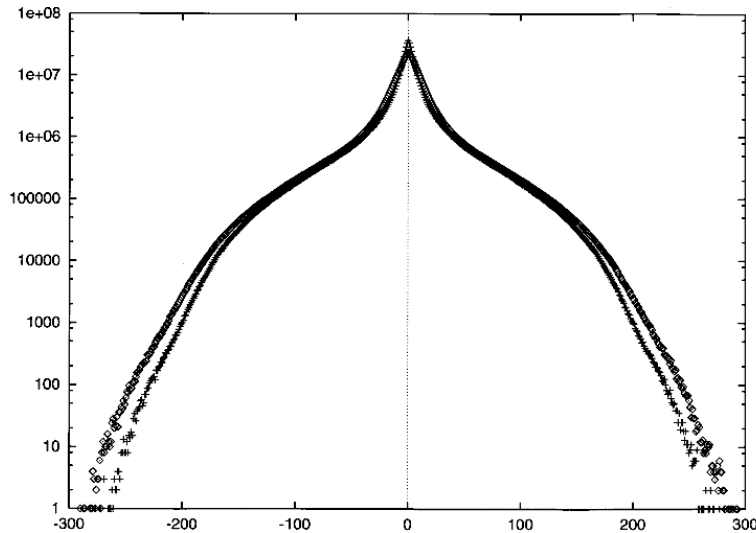
Említettük, hogy amennyiben az aktivitás kicsi, a hozam eloszlása hasonló lesz a csoportméret eloszláshoz. Az aktivitás az egy adott pillanatban üzletelő csoportok számát jelentette a fenti mikroszkopikus modellekben. Ha azonban feltételezzük, hogy egy adott időpillanatban egyszerre csak egy üzlet bonyolítható le (a valóságban ez történik), akkor az aktivitás azt fogja jelenteni, hogy az üzletkötések által kialakított árváltozás milyen időintervallumban jött létre, tehát az aktivitás mértéke felfogható, mint a hozamhoz tartozó időkülönbség,  $\Delta t$ , nagysága.



2.15. ábra. Átmenet a Gauss eloszlásba a perkolációs mikroszkopikus modell esetén. Az ábrán a hozam sűrűségfüggvényét láthatjuk különböző aktivitás értékek mellett. A aktivitás értéke a csoportok teljes számának százalékában kifejezve 1.25 %-tól ( $\diamond$ ), 20 %-ig ( $*$ ) változnak [Stauffer and Penna, 1998].

A korábbi gondolatmenet értelmében akkor kapunk hatványfüggvény lecsengésű eloszlást, ha az aktivitás alacsony, azaz  $\Delta t$  kicsi, és ugyanez volt a feltétel empirikus hozamok vizsgálatánál is. Ha  $\Delta t$  nő, az eloszlás skálázása megszűnik (az eloszlás már nem írható le egy állandó exponensű Lévy eloszlással), és tovább növelve az időkülönbséget végül az eloszlásunk Gauss eloszlású lesz. Ezt a viselkedést vizsgálták meg a perkolációs mikroszkopikus modell esetén [Stauffer and Penna, 1998]. Amennyiben az aktív csoportok száma nem haladja meg az összes csoport számának 0.2%-át, a

különböző aktivitáshoz (tehát különböző időkülönbséghez) tartozó hozam eloszlások egymásba skálázhatóak. Ha azonban az aktivitás 1.25 %-nál nagyobb, a hozam eloszlása már nem mutat hatványfüggvény lecsengést. Tovább növelve az aktivitást átmenet figyelhető meg a Gauss eloszlásba, ld. 2.15 ábra.



2.16. ábra. Perkolációs modell hozam eloszlása a különböző aktivitási modellek esetében. ( $\diamond$ ) konstans aktivitás, azaz  $\epsilon = 0$ ; (+) független az ártól,  $\epsilon = 0$ . [Chang and Stauffer, 1999]

Az eddig említett modellekben az aktivitás független volt a részvény áratól, időben állandó érték volt. A valóságban azonban az aktivitás erősen függhet a részvény árának aktuális értékétől, attól hogy mennyire tér el a fundamentális ártól. Az aktivitásnak ez a viselkedése könnyen beépíthető a perkolációs modellbe [Chang and Stauffer, 1999]. A vásárlás illetve az eladás valószínűsége,  $a$ , ne legyen egyenlő, hanem, ha az ár túl magas, több csoport akarjon eladni, mint venni, illetve fordítva: ha túl alacsony, a csoportok nagyobb valószínűséggel akarjanak venni, mint eladni. Létezik tehát egy visszatérítő erő, amely az árat a fundamentális értékébe akarja visszavinni. Jelöljük  $P_0$ -val a fundamentális árat (amely az egyszerűség kedvéért legyen konstans a folyamat során), és a részvény árának ettől való eltérését jelölje  $\Pi(t) = \log[P(t)/P_0]$ . A vásárlás,  $\pi_b$ , illetve az eladás,  $\pi_s$ , valószínűsége:

$$\begin{aligned}\pi_s &= (1 + \epsilon\Pi)a \\ \pi_b &= (1 - \epsilon\Pi)a,\end{aligned}\tag{2.42}$$

ahol  $a$  a korábbi modellben szereplő aktivitási hányad, és  $\epsilon$  egy konstans szám, ami annak az erősségét jelzi, hogy mennyire függ a vásárlás illetve eladás aktivitásának különbsége az ár fundamentális értéktől való eltérésétől. Nyilván  $\epsilon = 0$  esetén az eredeti modellt kapjuk vissza.

A hozam ezúttal is legyen arányos a vásárló illetve eladó csoportok méreteinek különbségével:

$$\log[P(t)/P(t-1)] = \Pi(t) - \Pi(t-1) = \sum_{s=1}^{\infty} (N_s^+ - N_s^-) \frac{s}{N}, \quad (2.43)$$

ahol  $N_s^+$  és  $N_s^-$  az  $s$  méretű vásárló illetve eladó csoportok száma.

A fenti modellt szimulálva az a meglepő eredmény adódott, hogy noha  $\Pi(t)$  időbeli változása jelentősen megváltozik, ha  $\epsilon \neq 0$ , de a logaritmikus hozam eloszlásában illetve az ár abszolút értékének autokorrelációjában nincs jelentős változás.

## 3. fejezet

# Tőzsdeindex fluktuációjának vizsgálata

Ez a fejezet a tőzsde index (vagy egyetlen részvény) időbeli fluktuációjával foglalkozik. Először egy empirikus vizsgálatot tekintünk, amely a 2.1.3 fejezethez hasonlóan az index fluktuáció eloszlásának skálázásával foglalkozik. Azt vizsgáljuk, hogy a skálázás érvényessége, illetve a normális eloszlástól való távolság hogyan függ a logaritmikus árfluktuációhoz tartozó időkülönbségtől.

A 3.2 fejezet egy egyrészvényes, sok résztvevős mikroszkopikus modellt elemez. Azt vizsgáljuk, hogy ez a modell mennyire képes reprodukálni a 2.1 fejezetben említett, árfluktuációra tett empirikus megfigyeléseket, elsősorban az eloszlás leptokurtikus mivoltát, illetve az árfluktuáció abszolút értékének lassú lecsengésű autokorrelációját. Megvizsgáljuk továbbá, hogy teljesül-e a skálázás, és ha igen, milyen határok közt, illetve ebből a szempontból összehasonlítjuk a modellt egy másik, perkolációs rácson szimulált modellel [Stauffer and Penna, 1998].

### 3.1. Eloszlás skálázásának határa

A 2.1.3 fejezetben azt láttuk, hogy a különböző  $\Delta t$  időkülönbségekhez tartozó logaritmikus árfolyam különbségek eloszlásai egymásba skálázhatók, (2.9 ábra), azaz ugyanazon eloszlással reprezentálhatóak. Ez az eloszlás leptokurtikus, tehát a Gauss eloszlásnál lassabb lecsengést mutat. Azt is láttuk azonban, hogy az eloszlás véges szórással rendelkezik, ezért ha az összegzendő változók száma a végtelenhez tart, azaz ha a logaritmikus árváltozáshoz tartozó időkülönbség nagyon nagy, akkor a centrális határeloszlás tétel miatt a határeloszlás Gauss lesz. Ebből következik, hogy létezik egy karakterisztikus időkülönbség  $\tau_S$ , amelynél kisebb időkülönbségekre a logaritmikus árváltozások eloszlása skálázható, a közös eloszlás leptokurtikus. Nagyobb időkülönbségekre azonban a skálázás már nem érvényes.

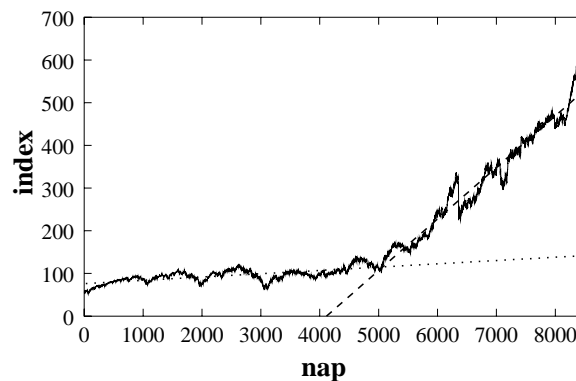
Ahhoz, hogy meghatározzuk a karakterisztikus időkülönbséget, abból a felte-

vésből indulunk ki, hogy a logaritmikus árváltozás eloszlása egy bizonyos tartományban ( $r \leq 6\sigma$ ), jól jellemezhető egy szimmetrikus, stabilis Lévy eloszlással [Mantegna and Stanley, 1995]. A rendelkezésre álló adatokból számolt hisztogramra a különböző időkülönbségek esetén Lévy illetve Gauss eloszlásokat illesztettünk.

Mértük egyrészt, hogy a legjobban illeszkedő Lévy eloszlás exponense hogyan változik az időkülönbség függvényében. Ahhoz ugyanis hogy a skálázás érvényes maradjon, az exponensnek állandó értéken kell maradnia. Másrészt azt vizsgáltuk, hogy az illesztett Lévy illetve Gauss eloszlás hibája hogyan viszonyul egymáshoz az időkülönbség függvényében.

### Felhasznált adatok

Az analízishez a napi felbontású – minden egyes napra csak az index záróértéke van megadva – S&P500 tőzsdeindexet használtuk az 1962 július 3 és 1995 december 29 közötti időtartományban, ami 8431 üzleti napnak felel meg, 3.1 ábra. Az ábráról világosan látszik, hogy az 1980-as évek elejétől megváltozik a trend. Míg az első 5000 napban az index várhatóértékének emelkedése lassú, addig az ezt követő időszakban egy sokkal drasztikusabb emelkedés látható. Ezen kívül a második időszakban a fluktuációk mértéke is nagyobb. A fenti okok miatt a skálázás vizsgálatánál külön-külön is vizsgáltuk ezt a két időtartományt.



3.1. ábra. A Standard and Poor's 500 New York-i tőzsdeindex időbeli fejlődése lineáris skálán. A pontozott, illetve a szaggatott vonalak az első 5000 illetve a második körülbelül 3400 üzleti nap index értékeire illesztett egyenesek, amelyek világosan mutatják a két időtartomány trendjei közti különbséget. (Az eltérés az 2.1/a) ábrához képest, hogy ott egy kicsit hosszabb időtartomány látszik, azonkívül ott az indexértékek logaritmikus skálán vannak ábrázolva.)

### Az árfluktuáció megfelelő mértéke

A dolgozat első fejezetében említettük, hogy a tőzsdei ár fluktuációját több, különböző mennyiség mérésével lehet vizsgálni, mint például a növekmény,  $z_{\Delta t}(t) = (t + \Delta t) -$

$p(t)$ , vagy a logaritmus növekmény (hozam)  $r_{\Delta t}(t) = \log [p(t + \Delta t)] - \log [p(t)]$ . A gazdasági folyamatok vizsgálatánál általában az utóbbit szokták előnyben részesíteni, mert ez valamennyire kiszűri az index időbeli fejlődésében fellépő trendet. Ahhoz, hogy megmutassuk, hogy akár egy Gauss folyamat is eredményezhet leptokurtikus eloszlást, ha a növekményt használjuk a logaritmus hozam helyett, egy mesterséges adatsort generáltunk. Legyen az index időbeli fejlődése az alábbi sztochasztikus folyamat:

$$x(t) = [x(t-1) + \xi(t)]b, \quad (3.1)$$

ahol a  $\xi(t)$  Gauss eloszlással rendelkezik:

$$P(\xi) = \frac{1}{\sqrt{2\pi\sigma^2(t)}} \exp \left\{ -\frac{\xi^2}{2\sigma^2(t)} \right\}, \quad (3.2)$$

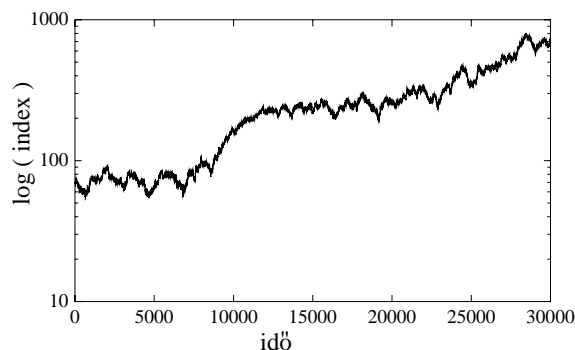
amelynek szórása időfüggő:

$$\sigma(t) = \sigma(t-1)b. \quad (3.3)$$

A kezdeti értékeket,  $\sigma(t=0)$  és  $x(t=0)$  valamint a  $b$  paramétert úgy választottuk, hogy megfeleljenek az empirikus értékeknek, amit az általunk felhasznált S&P500 index adatokból becsültünk meg:

$$b - 1 = 7.68 \cdot 10^{-5}, \quad \sigma(t=0) = 0.4, \quad x(t=0) = 70.$$

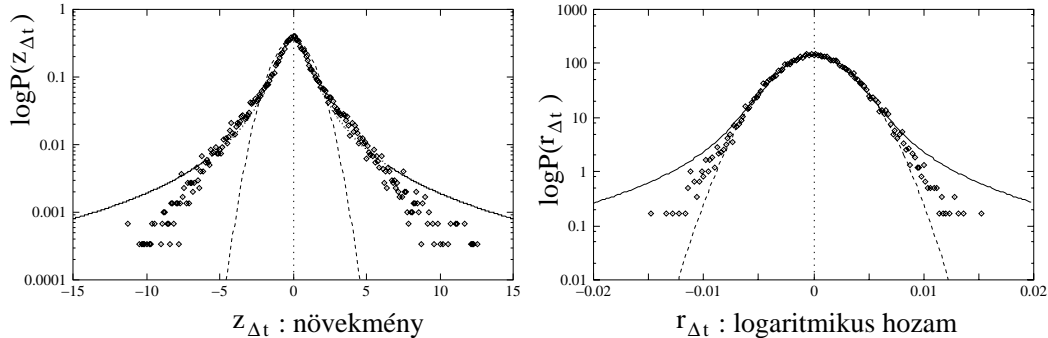
Az így definiált sztochasztikus folyamattal 30000 időlépést generáltunk, ami a durván 10000 nappól álló empirikus időszorral összevetve napi 3 indexértéknek felel meg. Az empirikus (3.1 ábra) és a mesterséges (3.2 ábra) idősor – legalábbis szemre – elég hasonlóknak tűnnek. A legszembetűnőbb különbség, hogy a mesterséges folyamatban nem léteznek akkora ugrások, mint például az S&P500 indexben az 1987-es októberi „fekete hétfő”. Ezek után vizsgáljuk meg a mesterséges idősoron, hogy milyen kü-



3.2. ábra. Szimulált tőzsde-index idősor. A sztochasztikus folyamat egy „infláló” véletlen bolyongásnak felel meg, amit a (3.1) egyenlet határoz meg.

lönbséget okoz ha az árfluktuáció mérésére a két különböző, fent említett mértéket

használjuk. A 3.3 ábra szemlélteti a két mérték esetén a valószínűsűrsűrűség függvényét. Annak ellenére, hogy a mesterséges idősorban a zaj Gauss eloszlású, ha a folyamat változását a növekménnyel,  $z_{\Delta t}$ , írjuk le, az eloszlás mégsem Gauss lesz, hanem egy ennél lassabban lecsengő függvény. Ennek az az oka, hogy a szimulált idősor egy időfüggő várhatóértékkel, illetve szórással rendelkezik, ami a valós adatokon is megfigyelhető. Azonban ha a fluktuációk leírására a logaritmikus hozamot használjuk, azt láthatjuk, hogy a Gauss eloszlás jól jellemzi a fluktuációk eloszlását. Ez a példa azt illusztrálja, hogy a vizsgálatoknál előnyösebb a logaritmikus hozamot használni az árváltozások fluktuációjának leírására.



3.3. ábra. A mesterséges idősor növekményeinek eloszlása a két különbözően definiált árváltozás esetén. A bal oldalon a növekmény,  $z_{\Delta t}$ , a jobb oldalon a logaritmikus hozam,  $r_{\Delta t}$ , eloszlása látható. Mindkét eloszlásra illesztettünk egy Lévy és egy Gauss eloszlást. Világosan látszik, hogy az egyszerű növekményt véve a változás mértékének az eloszlás a Gaussnál lassabban lecsengő, leptokurtikus függvény lesz.

### 3.1.1. Illesztési eljárás

A célunk az, hogy megvizsgáljuk, hogy az árváltozások eloszlásának skálázása az árváltozáshoz tartozó időkülönbség,  $\Delta t$  függvényében teljesül-e, és ha igen, milyen határokon belül. Ebből a célból az empirikus adatokból készített hisztogramra Gauss illetve szimmetrikus Lévy eloszlást illesztünk, és azt vizsgáljuk, hogy a különböző időkülönbségek esetén az illesztett Lévy eloszlás exponense változik-e, illetve, hogy az illesztett függvények hibái milyen arányban állnak egymással.

Míg a normális eloszlás illesztése egyszerűbb feladat, addig a Lévy eloszlás illesztése kis nehézséget okoz, lévén hogy csak a karakterisztikus függvényének (sűrűségfüggvény Fourier transzformáltjának) létezik analitikus alakja. Így az illesztést a Fourier térben kell elvégezni, majd az illesztett görbét visszatranszformálni a valós térbe.

Először elkészítjük a tapasztalati sűrűségfüggvényt (hisztogramot), azaz megnézzük, hogy mennyi a  $\Delta x$  intervallumba eső pontok száma,  $N_x$ , mivel  $N_x/N \simeq P(x <$

$\xi \leq x + \Delta x) \simeq p(x)\Delta x$ , ahol  $N$  az összes rendelkezésre álló pontok száma. (Sajnos a rendelkezésre álló adatok kis száma miatt – összesen  $N = 8431$  értékünk van – a hisztogram elkészítéséhez a rendelkezésre álló statisztika nem a legjobb.) Fontos az intervallumméret  $\Delta x$  helyes megválasztása. Túl nagy intervallumméret esetén a hisztogramot leíró pontok száma kevés lesz, amely a Fourier transzformáció pontatlanságát vonja maga után. Túl kicsi intervallumméret esetén azonban a hisztogram simaságán esik csorba, azaz az egymást követő intervallumokra eső pontok száma nagy szórást mutathat, ami megint csak a Fourier transzformáció pontatlanságához vezet. Az analízisünkben akkora intervallumot választottunk, hogy az ugyanazon intervallumba eső pontok száma átlagosan az összes rendelkezésre álló pontok számának,  $N$ , a gyöke legyen. Ezután a hisztogramban fellépő szórást egy simító algoritmussal (spline interpoláció) redukáltuk.

A spline interpoláció a  $\xi_i$  mért adatokra illesztett regressziós görbe,  $g(x)$ , amely az  $x_i$  mérési pontokban másodrendben folytonos parabolaívvekből áll. A harmadik derivált lehet különböző értékű a kapcsolódási pont két oldalán. Späth javaslata szerint [Späth, 1978] ez az  $r_i$  különbség legyen arányos az alappont ordináta,  $\xi_i$ , és a spline ordináta,  $g_i$ , különbségével :

$$r_i = p_i(\xi_i - g_i), \quad (3.4)$$

ahol  $p_i$  a simítási paraméter, amely pontról pontra változhat. Minél nagyobb  $p_i$  értéke, a regressziós görbe annál jobban hozzásimul a mért adatokhoz. Ezért az origó közelében nagy  $p_i$  értéket választottunk, hiszen itt fontos a pontos illeszkedés; az origótól távolodva pedig egyre csökkentettük, hogy az itt fellépő nagy szórást kisimítsa:

$$p_i = \frac{p_0}{x_i^2 + \Delta p}, \quad (3.5)$$

ahol  $x$  az abszcissa értéke.

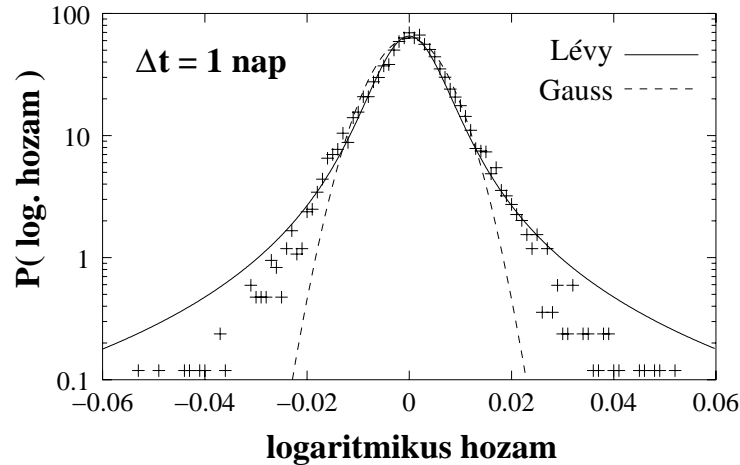
A simítás után a kisimított függvényt diszkrét Fourier transzformációval a Fourier térbe transzformáljuk, és itt illesztjük rá a szimmetrikus Lévy eloszlás karakterisztikus függvényét, amelyet a (2.6) egyenlet definiál. Az illesztést a legkisebb négyzetek módszerével végeztük, ami azt jelenti, hogy az alábbi kétváltozós függvény minimumát kell meghatározni:

$$\varepsilon^2(\alpha, a_\alpha) = \sum_i \left[ \tilde{p}(z_i) - \exp \{-a_\alpha z_i^\alpha\} \right]^2, \quad (3.6)$$

ahol  $\tilde{p}(z)$ , a simított hisztogram Fourier transzformált függvénye.

A kiszámolt  $\alpha$  illetve  $a_\alpha$  paraméterű Lévy eloszlás karakterisztikus függvényét inverz Fourier transzformáltuk – ennek eredménye,  $L_\alpha(x)$  –, hogy kiszámolhassuk, mekkora eltérést mutat a tapasztalati sűrűségfüggvényhez képest,  $\varepsilon_L^2 = \sum_i |p(x_i) - L_\alpha(x_i)|^2$ . Majd a hisztogramunkra illesztettünk egy normális eloszlást is,  $N_\sigma(x)$ , és

ennek is mértük az empirikus eloszlástól való négyzetes eltérését:  $\varepsilon_G^2 = \sum_i |p(x_i) - N_\sigma(x_i)|^2$ . Az illesztési folyamat eredményét a 3.4 ábra szemlélteti. Az ábráról azt olvashatjuk le, hogy noha a Lévy eloszlás felülbecsüli a tapasztalati sűrűségfüggvényt, ha a logaritmusos hozam  $r > 0.02$ , azonban a Gauss eloszlás legalább annyira alulbecsüli azt.



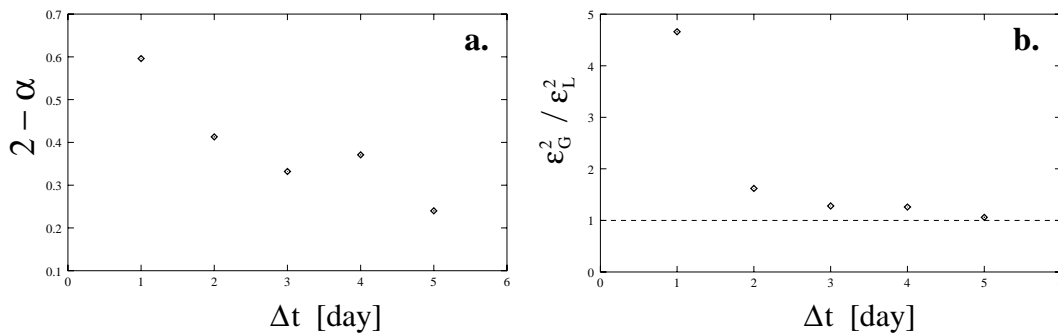
3.4. ábra. Az S&P500 index adatok logaritmusos hozamának hisztogramja (+-szal jelölve), illetve az illesztett Lévy (folytonos vonal) és Gauss (szaggatott vonal) eloszlások. A Lévy exponens értéke  $\alpha = 1.4 \pm 0.03$ , ami megközelítőleg megegyezik a Mantegna és Stanley [Mantegna and Stanley, 1995] által kapott exponens értékkel.

### 3.1.2. Karakterisztikus idők meghatározása

A skálázás feltétele, hogy az exponens,  $\alpha$  értéke ne változzon a hozam időkülönbségének,  $\Delta t$  függvényében. A karakterisztikus idő meghatározásának tehát legnyilvánvalóbb módszere, hogy az előző fejezetben említett illesztési eljárással mérjük az  $\alpha$  exponenst, mint  $\Delta t$  függvényét. Ennek az eredményét mutatja a 3.5/a) ábra. Azt láthatjuk, hogy az exponens értéke változik  $\Delta t$  függvényében már  $\Delta t = 2$  nap esetén is. Az is látszik azonban az ábráról, hogy – mivel az adatok száma csekély (8000 napi hozam) – az illesztési eljárás hibája nem elhanyagolható, amit a 4. napnál mutatkozó exponens érték növekedés mutat. Ezt leszámítva azonban egy világos trend figyelhető meg az  $\alpha(\Delta t)$  függvényben. Mivel a 2.1.3 fejezetben azt láttuk, hogy a skálázás érvényes a nagy-felbontású adatok esetén, azaz a hozam egy napnál kisebb időkülönbsége esetén [Mantegna and Stanley, 1995], nyilvánvalóan következik, hogy az általunk keresett karakterisztikus idő az egy nap nagyságrendjébe esik.

Mivel, mint említettük, az illesztési eljárásnak van egy bizonytalansága, nem elégedtünk meg csupán  $\alpha(\Delta t)$  függvény vizsgálatával, hanem azt is megnéztük, hogy az

illesztett függvényeknek a hisztogramtól való négyzetes eltérése (például (3.6) egyenlet) hogyan aránylik egymáshoz, és ez az arány hogyan változik, ha növeljük az időkülönbséget. A 3.5/b) ábra ezt a vizsgálatot szemlélteti. Azt láthatjuk, hogy  $\Delta t$  növelésével a hibák aránya csökken (a csökkenés trendje világosan látszik), és  $\Delta t = 5$  nap esetén, mindkét eloszlás nagyjából ugyanakkora hibával illeszkedik. Ez is alátámasztja az állításunkat, hogy néhány nap időkülönbség esetén a skálázás már nem érvényes.



3.5. ábra. Átmenet a Lévy eloszlásból Gauss eloszlásba. (a) Az illesztett eloszlás exponensének a Gauss eloszlás exponens értékétől való eltérése ( $2 - \alpha$ ) a hozamhoz tartozó  $\Delta t$  különbség függvényében. (b) Az illesztett Gauss és Lévy eloszlások négyzetes hibájának hányadosa  $\Delta t$  függvényében.

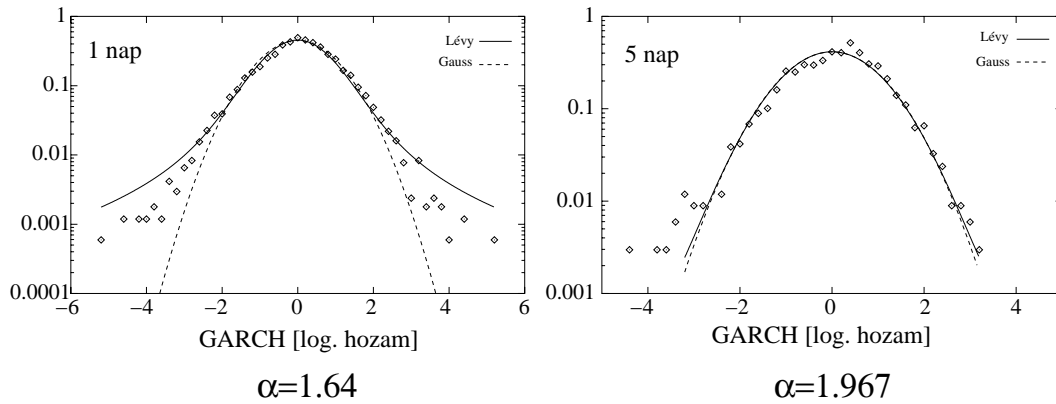
A 3.1 fejezet második alfejezetében azt taglaltuk, hogy mennyire fontos, hogy az index fluktuációinak leírásához a helyes mértéket válasszuk, ezzel kiszűrve olyan hatásokat, amelyek leptokurtikussá teszik az eloszlásfüggvényt (például az átlag időbeli változása). Ezért választottuk az időkülönbség helyett a logaritmikussá hozamot. Léteznek azonban sokkal kifinomultabb módszerek ezen effektusok kiszűrésére, ilyenek például a 2.1.4 fejezetben bemutatott ARCH, GARCH módszerek. Ha ezeket alkalmazzuk a nyers adatsorra, akkor a 3.1 táblázatból leolvashatjuk, hogy hogyan módosul az exponens értéke. Azt láthatjuk, hogy míg az ARCH módszer nem okozott jelentős változást, a GARCH módszerrel kapott exponens értéke sokkal nagyobb lesz, de még mindig nem éri el a normális eloszlásnak megfelelő exponens értéket ( $\alpha = 2$ ). Tehát továbbra is leptokurtikus lesz az eloszlás. Ez azt jelenti, hogy az eloszlás lassan lecsengő jellege nem csupán az időben változó várhatóérték és szórás miatt adódik.

3.1. táblázat. Az illesztett  $\alpha$  exponens értéke a különböző „szűrési eljárások” esetében.

Változó típusa	Exponens ( $\alpha$ )
Log. hozam	1.40
ARCH	1.44
GARCH	1.64

Ha a GARCH módszerrel szűrt logaritmikussá hozam eloszlását és az erre illesztett

Gauss és Lévy függvényeket vizsgáljuk az időkülönbség függvényében (3.6 ábra) azt láthatjuk, hogy  $\Delta t = 5$  nap esetén az eloszlás szinte teljesen Gauss jelleggel bír; a Lévy eloszlás exponense  $\alpha = 1.967$ , a normális eloszlásával egyezik meg.



3.6. ábra. Az empirikus eloszlás konvergenciája Lévy típusú viselkedésből Gauss eloszlásba, a GARCH módszerrel kezelt logaritmikus hozam esetén. (a)  $\Delta t = 1$  nap, az illetett Lévy eloszlás exponense  $\alpha = 1.64$ , jóval meghaladja a GARCH nélkül számolt eloszlás exponensét ( $\alpha \simeq 1.4$ ) (b)  $dt = 5$  nap, az exponens értéke  $\alpha = 1.967$ , ami azt jelenti, hogy GARCH módszer esetén az eloszlás már 5 nap után jó közelítéssel Gaussnak tekinthető.

### Gauss eloszláshoz való konvergencia

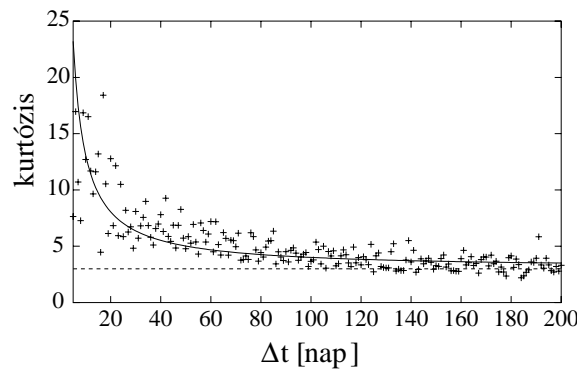
Az előző fejezet azt mutatta, hogy amennyiben a logaritmikus hozam időkülönbsége meghaladja az egy napot, a hozam eloszlásának skálázása nem érvényes. Tehát definiálhatunk egy karakterisztikus időt,  $\tau_S$ , ami a skálázás érvényességének határát jelöli. Abból a tényből, hogy a skálázás nem érvényes, azonban nem következik, hogy az eloszlás ezen a skálán Gauss lesz. A 3.5 ábrán ugyan azt láthatjuk, hogy az illetett Gauss és Lévy sűrűség-függvények hibáinak (a hisztogramtól való négyzetes eltéréseik) hányadosa tart az egyhez, ez azonban csak annyit jelent, hogy a Lévy eloszlás nem illeszkedik jobban mint a Gauss eloszlás. A skálázás érvényességének megszűnését jellemző  $\tau_S$  karakterisztikus időn kívül definiálhatunk egy másikat,  $\tau_G$ , amely arra az időskálára jellemző, amelynél az index-hozam eloszlása Gauss eloszlású lesz. Ez utóbbit például a kurtózis (negyedik centrális momentum) konvergenciájával definiálhatjuk [Pasquini and Serva, 2000]. Tegyük fel, hogy a logaritmikus hozam eloszlása a (2.9) egyenletben leírt, exponenciálisan levágott Lévy eloszlással jellemezhető. Ebben az esetben az ilyen eloszlással rendelkező változók összegének kurtózisa analitikusan számolható:

$$\kappa(N) = \frac{M_4}{M_2^2} = 3 + \frac{1}{N} \left[ -\frac{\cos\left(\frac{\pi\alpha}{2}\right)}{a_\alpha} \frac{(\alpha-2)(\alpha-3)}{\alpha(\alpha-1)} \mu^{-\alpha} \right], \quad (3.7)$$

ahol  $N$  az összegzett változók száma, és  $\mu$  a *levágási paraméter*, amely fordítottan arányos a levágási hosszal. Ez egy olyan karakterisztikus érték, amely azt a nagyságrendet jellemzi, amelynél nagyobb változók esetén az eloszlás már nem hatványfüggvényt, hanem exponenciális függvényt követ.

Ahhoz, hogy a változók összegének kurtózisát analitikusan számolhassuk, feltételeztük, hogy az eloszlás exponenciális levágású. Ugyanakkor a bevezetőben említettük, hogy a legújabb eredmények arról számolnak be, hogy a levágás is hatványfüggvény lecsengést mutat, csak nagyobb exponens értékkel. Úgy gondoljuk, hogy ez a különbség nem változtatja az eredményt lényegesen, és a változás valószínűleg csak növeli a  $\tau_G$  karakterisztikus időt.

A (3.7) egyenletben szereplő  $(\alpha, a_\alpha, \mu)$  paramétereket úgy kaphatjuk meg, ha  $N = 1$  esetén illesztünk egy levágott Lévy eloszlást az empirikus adatokra. Ha az így meghatározott paramétereket behelyettesítjük (3.7)-be, akkor azt láthatjuk (3.7 ábra), hogy az így meghatározott kurtózis, mint a változók összegének a függvénye,  $\kappa(N = \Delta t)$ , jól illeszkedik az empirikusan számolt kurtózis értékekre. Továbbá azt olvashatjuk le az ábráról, hogy az az időkülönbség, amelynél a kurtózis értéke megközelíti a normális eloszlásnak megfelelő kurtózis értéket ( $\kappa = 3$ ), néhány hónap nagyságrendjébe esik.



3.7. ábra. Az index-hozam kurtózisa, a hozamhoz tartozó időkülönbség függvényében, illetve a kurtózis analitikusan számított értéke. A paraméterek értékeit a hozam eloszlására illesztett exponenciálisan levágott Lévy eloszlás paramétereiből határoztuk meg:  $\alpha = 1.4$ ,  $a_\alpha = 8.6$ ,  $\mu = 0.009$

Összefoglalva az eredményeinket azt láthatjuk, hogy legalább két különböző karakterisztikus időt definiálhatunk. Az egyik a hozam eloszlásának skálázására jellemző,  $\tau_S$ , ennek az értéke egy nap nagyságrendjébe esik, míg a másik karakterisztikus idő,  $\tau_G$  azt az időkülönbséget jellemzi, amelynél a hozam eloszlása ténylegesen Gauss eloszlású lesz. Ez utóbbi az egy hónap nagyságrendjébe esik, tehát  $\tau_S \ll \tau_G$ .

## 3.2. Egyrészesvényes mikroszkopikus modell

Ebben a fejezetben egy mikroszkopikus modellel foglalkozunk [Kullmann and Kertész, 2001c], amely a 2.3.1 fejezetben említett „csordaszellemű” modellen [Cont and Bouchaud, 2000] alapszik. A tőzsdei ügynökök különböző méretű csoportokat alkotnak, és a csoporton belül minden résztvevő ugyanazt a stratégiát követi, vásárol, elad, vagy nem üzletel,  $\phi_i(t) \in \{-1, 0, 1\}$ . A részvény árának változása arányos a vásárlási és az eladási szándékkal rendelkező ügynökök teljes számának a különbségével:

$$r(t) \sim \sum_{i=1}^n s_i \phi_i(t),$$

ahol  $\phi_i(t)$  az  $i$ -edik csoport stratégiája, és a csoportok egymástól függetlenül hozzák meg döntéseiket.

A csoportok kialakulása a preferált növekedési modellnek megfelelően történik, amit a következő fejezetben tárgyalunk. A 3.2.2 fejezetben ennek a kialakulási folyamatnak a leírásával foglalkozunk. Jelölje  $\mathcal{P}_i(s, t)$  annak a valószínűségét, hogy a  $t$ -edik időpontban az  $i$ -edik csoport mérete  $s$ . Ennek a mennyiségnek az analitikus megoldását számoljuk ki. Erre azért van szükség, mert noha a modellünkben a csoportok mérete még az üzletelési folyamat előtt kialakul, és annak során nem változik, úgy gondoljuk, hogy valóságosabb lenne egy olyan gazdasági modell bevezetése, ahol ennek a két folyamatnak az időskálája nem válik szét.

Bevezettünk egy ár függő aktivitást, amely azonban eltér a 2.3.2 fejezetben leírtól, mert nemlineáris módon függ az árnak a fundamentális értéktől való eltérésétől. Ezt bővebben a 3.2.3 fejezet tárgyalja.

### 3.2.1. Csoportok kialakításának modellje

Láttuk, hogy amennyiben a csoportok kialakulása egy perkolációs modellel történik, a csoportméret eloszlás hatványfüggvény lecsengésű lesz. (Feltételezve, hogy vagy a perkolációs kritikus pontban vagyunk, vagy átlagoljuk a csoportokat a különböző betöltési valószínűségek esetén).

Egy másik csoport kialakítási lehetőség, amely szintén hatványfüggvény lecsengésű csoport eloszlást eredményez, az úgynevezett *preferált növekedési modell* [Simon, 1955]. (Természetesen nagyon sokféle módon lehetne hatványfüggvény eloszlású csoportokat kialakítani, itt azonban fontos szempontnak tartjuk, hogy a kialakulási mechanizmus valamelyest valóságos legyen, tehát olyan módszerrel írjuk le a csoportok kialakulását, amely a valóságban is elképzelhető.)

A preferált növekedési modellt először hálózatok kialakulásának leírására vezették be. Képzelnünk el egy dinamikusan fejlődő hálózatot (összekötött csomópontok

halmazát), ahol minden időlépésben új csomópontok csatlakoznak a már meglévőkhöz. Minél több kapcsolattal rendelkezik egy, már meglévő csomópont, annál nagyobb valószínűséggel fog ehhez kapcsolódni az új elem.

Az utóbbi időkben kimutatták, hogy ilyen modellek alkalmasak olyan látszólag különböző hálózatok leírására, mint például a WWW hálózat [Albert et al., 1999], [Barabási and Albert, 1999], az Internet [Faloutsos et al., 1999], a tudományos hivatkozások statisztikája [Redner, 1998], stb. Ezek a rendszerek azzal a közös tulajdonsággal rendelkeznek, hogy az egy csomópontból kiinduló kötések száma (azaz egy csomópont kapcsolatainak száma) hatványfüggvény lecsengésű, tehát skálainvariáns. Kiderült, hogy ez a skálainvariancia a rendszer kialakulási folyamatának két fontos tulajdonságából adódik: egyrészt, hogy a rendszer az *időben folyamatosan növekszik*, valamint hogy egy új elemnek egy már létező csomóponttal való kapcsolódási valószínűsége *arányos* a csomópontokból kifutó kapcsolatok számával.

Tegyük fel, hogy az  $i$ -edik csomópont  $s_i$  kapcsolattal (kötés) rendelkezik, ekkor annak a valószínűsége, hogy az új elem  $i$ -edik csomóponttal csatlakozik,  $\Pi_i$ , arányos  $s_i$ -vel:

$$\Pi_i = \frac{s_i}{\sum_{i=1}^n s_i}, \quad (3.8)$$

ahol a nevezőben szereplő összeg a rendszerben szereplő összes kötés száma, és  $n$  a már létező csomópontok száma. Az így kialakuló hálózatban egy csomópontból kiinduló kötések számának valószínűsége függvénye hatványfüggvény lecsengésű lesz [Albert et al., 1999]:

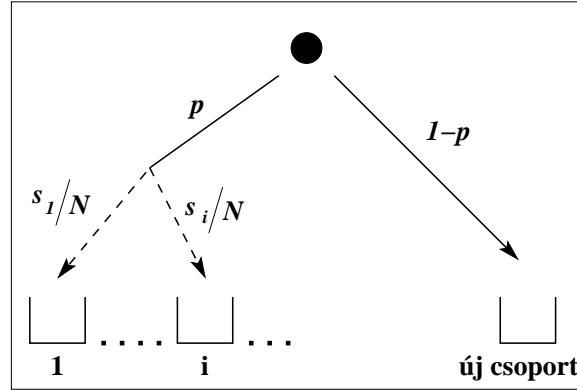
$$\mathbf{P}(s) \underset{s \rightarrow \infty}{\sim} s^{-3}, \quad (3.9)$$

Ahhoz, hogy ezt a modellt alkalmazhassuk az általunk vizsgált gazdasági folyamathoz, kis módosítás szükséges, ugyanis nekünk független csoportok kialakulására van szükségünk, míg a hálózati modell egy összefüggő halmazt ír le, azaz minden csomópontból bármely másikba eljuthatunk kötések keresztül. A módosított preferált növekedési modellt a 3.8 ábra szemlélteti. Kezdetben,  $t = 0$ , egyetlen csoport létezik, egyetlen elemmel. Minden időpillanatban egy új elem érkezik a rendszerbe, amely  $p$  valószínűséggel egy, már meglévő csoporthoz csatlakozik, és  $q = 1 - p$  valószínűséggel egy új csoportot alakít ki. Annak a valószínűsége, hogy a meglévő csoportok közül az  $i$ -edikhez kapcsolódjon, arányos az  $i$ -edik csoport méretével,  $\Pi_i \sim s_i$ . Az így kialakuló rendszerben a csoportméret eloszlás az alábbi alakot ölti:

$$\mathbf{P}(s) \underset{s \rightarrow \infty}{\sim} s^{-1-1/p}. \quad (3.10)$$

Amennyiben  $p = 0.5$  – tehát az új elem ugyanakkora valószínűséggel választja, hogy új csoportot alakítson-e ki, vagy már valamelyik létezőhöz kapcsolódjon –, a (3.10) egyenletben szereplő exponens értéke megegyezik a (3.9) egyenletben ismertetett eloszlás exponensével:  $(-1 - 1/p)|_{p=0.5} = -3$ .

Mint azt többször említettük, alacsony aktivitás esetén a csoportméret eloszlás megegyezik a hozam eloszlásával. Ez utóbbi exponense a legutóbbi vizsgálatok szerint  $\alpha \simeq 4$  [Gopikrishnan et al., 1998, Gopikrishnan et al., 1999]. Ahhoz tehát, hogy a modell által létrehozott hozam eloszlás aszimptotikus viselkedésének exponense megegyezzen a valós adatokon mért eredményekkel,  $p = 1/3$  választással kell élnünk.



3.8. ábra. Az ábra a „csordaszellemű” gazdasági modell csoportjainak kialakulását szemlélteti. Az ábra tetején látható pont az új elem, amely  $p$  valószínűséggel egy, már meglévő csoporthoz kapcsolódik, és  $q = 1 - p$  valószínűséggel egy új csoportot alakít ki.

### 3.2.2. A csoportméret eloszlás teljes megoldása

Ebben a fejezetben a fent bemutatott csoport kialakulási modellhez kapcsolódó különböző eloszlások analitikus megoldásait számoljuk ki. A számítás részleteit az (A) függelék taglalja, itt csak az eredményeket említjük meg.

A csoportméret eloszlás,  $\mathbf{P}(s, t)$  aszimptotikus viselkedését,  $\mathbf{P}(s \rightarrow \infty, t = \infty) = \mathbf{P}(s)$  a (3.10) egyenlet írja le. A teljes időfüggő megoldás tetszőleges,  $s$  csoportméret esetén az alábbi alakot ölti: (ld az A.4 fejezetet)

$$\mathbf{P}(s, t) = \underbrace{\sum_{l=1}^s (-1)^{l-1} \binom{s-1}{l-1} \left[ \frac{1-p}{1+lp} + \frac{p+lp}{1+lp} \frac{\Gamma(t-lp)}{\Gamma(t+1)\Gamma(1-lp)} \right]}_{\mathbf{P}(s, t = \infty)}. \quad (3.11)$$

Mivel a fenti kifejezésben a [...] -ben szereplő második tag  $t^{-1-lp}$  szerint cseng le, ezért  $t = \infty$  esetén visszakapjuk az (A.6) egyenletben meghatározott aszimptotikus értéket. Erre utal a képletben szereplő  $\mathbf{P}(s, t = \infty)$  kifejezés, amely ugyanakkor  $s \rightarrow \infty$  esetén a (3.10) alakot ölti, mint arra az A.1 fejezet utal.

Meghatároztuk továbbá az egyes csoportok teljes időfüggő méreteloszlását is [Kullmann and Kertész, 2001b] (a levezetést az A.2 fejezet részletezi):

$$\mathcal{P}_i(s, t) = \overbrace{\sum_{l=1}^s (-1)^{l-1} \binom{s-1}{l-1} \frac{\Gamma(t-lp)}{\Gamma(t)\Gamma(1-lp)}}^{\mathcal{P}_1(s, t)} \times \left[ \sum_{b=i}^t \frac{\Gamma(b)\Gamma(1-lp)}{\Gamma(b-lp)} \binom{b-2}{i-2} p^{b-i} (1-p)^{i-1} \right]. \quad (3.12)$$

A fenti kifejezés meglehetősen bonyolult, amennyiben további számításokra akarjuk használni (például gazdasági folyamat analitikusan kezelése) nehezen kezelhető. Azonban hasonlóan az átlagos csoportméret eloszláshoz, az aszimptotikus határesetben ( $t, s \rightarrow \infty$ ) ez is egyszerűbb formát ölt:

$$\lim_{t, i \rightarrow \infty} \mathcal{P}_i(s, t) = \left( \frac{i}{t} \right)^p. \quad (3.13)$$

A fent említett eloszlások közül néhányra már született korábban megoldás [Barabási et al., 1999, Simon, 1955, Dorogovtsev et al., 2000]. Ezek azonban egyrészt a hálózati modell esetén érvényesek, amely, noha csak kicsit, de eltér az általunk használt csoport kialakítási modelltől, másrészt nem tartalmazzák csoportok teljes időfüggő méreteloszlásának megoldását.

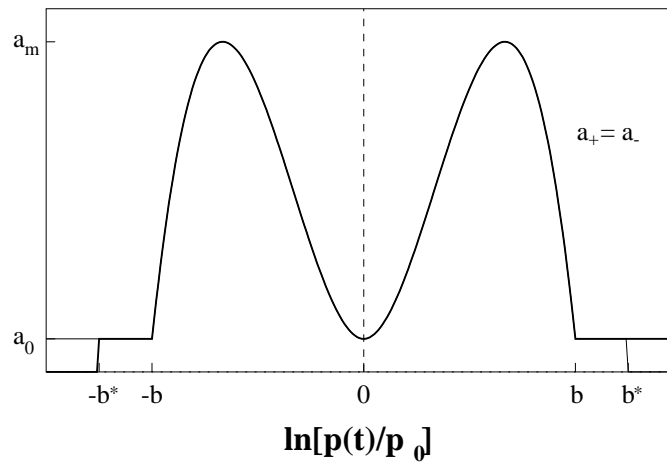
A gazdasági modellünkhöz szorosabban ugyan csak a csoportméret eloszlás, ennek is csak a stacionárius esete,  $\mathbf{P}(s)$  kapcsolódik, de mint azt korábban kiemeltük, lehetséges lenne egy olyan modell bevezetése, amelyben a csoport kialakulásának és az üzletkötés folyamatának időskálái nem válnak szét, sőt, a csoportok kialakulása esetleg függ az aktuális árfolyam értéktől. Ebben az esetben fontos lehet mind a teljes rendszer, mind az egyedi csoportok időfüggő méreteloszlásának ismerete.

### 3.2.3. Nemlineáris aktivitás

Ebben a fejezetben az aktivitás (egy adott pillanatban üzletelő csoportok hányada) ár-függésének kérdésével foglalkozunk. A 2.1.2 fejezetben említettük, hogy a hozam abszolút értékének hatványfüggvény szerint lecsengő autokorrelációja az adott  $\Delta t$  időintervallumban létrejött üzletkötések számának,  $N_{\Delta t}$ , lassan lecsengő autokorrelációja miatt van [Plerou et al., 2000]. Az  $N_{\Delta t}$  mennyiség a mi esetünkben megegyezik az aktív csoportok méretének összegével, azaz

$$N_{\Delta t} = S_a = \sum_{i=1}^{n_a} s_i, \quad (3.14)$$

ahol  $n_a$  az aktív csoportok száma. Definiáljuk az aktivitást,  $(n_a/n)$  mint az aktuális ár, és a fundamentális ár közti eltérés,  $\log[p(t)/p_0]$ , nemlineáris függvényét, amit



3.9. ábra. Az aktivitások árfüggése.  $a_+$  a vásárlási,  $a_-$  az eladási szándékú csoportok hányadát jelöli. Az aktivitás harmadrendű függvénye  $\log[p(t)/p_0]$ -nek, ahol  $p_0$  a fundamentális értéket jelöli. Amennyiben  $\ln[p(t)/p_0] < |b^*|$  az aktivitások  $a_+$  és  $a_-$  egyenlőek. Ha  $\ln[p(t)/p_0] > b^*$  akkor  $a_+ = 0$ ,  $a_- = a_0$  míg az ellenkező esetben:  $\ln[p(t)/p_0] < -b^* \Rightarrow a_- = 0$ ,  $a_+ = a_0$ . Az  $a_0$  mennyiség a csoportok számának reciprokával egyezik meg.

a 3.9 ábra mutat. A vásárlási aktivitás, (vásárló csoportok hányada,  $a_+$ ) illetve az eladási aktivitás, ( $a_-$ ) közt csak abban az esetben teszünk különbséget, ha az árak a fundamentális értéktől való eltérése meghalad egy küszöbértéket,  $b^*$ -ot. Ekkor attól függően, hogy felette van-e a fundamentális árnak, vagy alatta, a vásárlási, vagy az eladási aktivitás nullára csökken, míg a másik nem változik. Ez hasonlóan a 2.3.2 fejezetben említett módszerhez [Chang and Stauffer, 1999] egy visszatérítő erőt jelent, ami meggátolja, hogy a részvény ára túlságosan messze kerüljön a fundamentális értéktől.

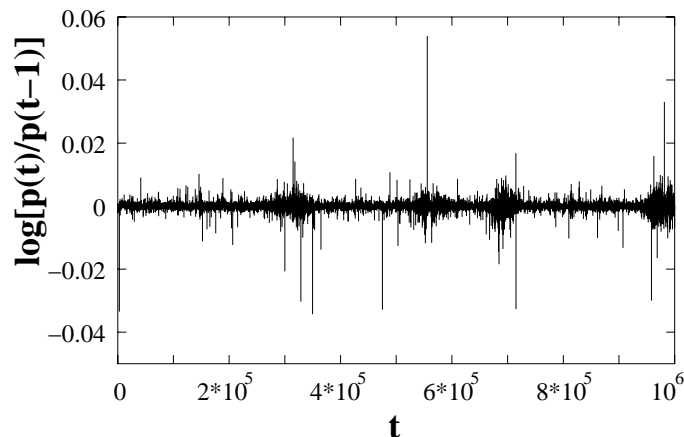
Az elsődleges célunk ezzel a definícióval azonban nem ez a visszatérítési erő bevezetése volt, hanem az, hogy biztosítsuk, hogy a folyamat egymást követő lépéseiben az aktivitások korreláltak legyenek. Adott aktivitási érték esetén, az ehhez kapcsolódó üzletkötések száma,  $N_{\Delta t}$  különböző lehet, de átlagosan nagyobb aktivitás nagyobb üzletkötési számot eredményez, ami azt jelenti, hogy ezek is korrelálni fognak az időben.

Az aktivitási függvény alakját intuitív módon határoztuk meg. Azon a hipotézisen alapul, hogy a fundamentális értéktől való kisebb eltérések inspirálják az üzletembereket, míg a nagy eltérések hatására az üzletemberek óvatosabbak lesznek, aktivitásuk visszaesik. Többféle függvényt is kipróbáltunk, és a tapasztalat azt mutatta, hogy az eredmények kevésbé érzékenyek a függvény típusának (másod vagy harmadrendű görbe) választásától.

### 3.2.4. Eredmények

A szimulációt a következőképpen végeztük. Először kialakítottuk a csoportokat a 3.2.1 fejezet alapján,  $p = 1/3$  paraméter értékkel. A rendszert  $N = 10^7$  méretűre választottuk, ami azt jelenti, hogy a csoportok száma körülbelül  $n = p \cdot N \simeq 3 \cdot 10^6$  lesz. Ezután meghatároztuk az aktivitás ár-függvényét, amit a 3.9 ábra szemléltet. Az  $a_0$  paraméter a csoportok számának reciproka. Az aktivitás maximum értékét,  $a_m = 500 \cdot a_0$  választottuk. A fundamentális ár értékét állandónak vettük a folyamat során, hiszen elsősorban a hozam és nem maga az ár időbeli függésére voltunk kíváncsiak, valamint a hozam korrelációjára. Ezután szimuláltuk a tőzsdei folyamatot.

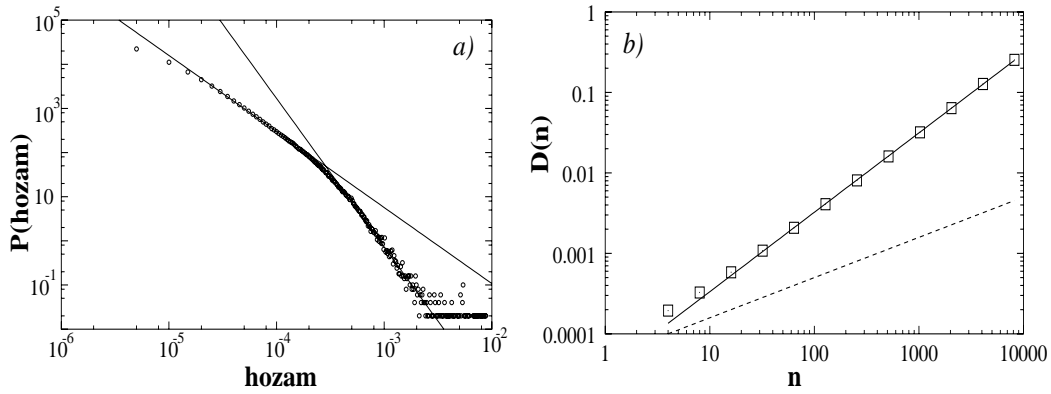
A 3.10 ábra egy példát mutat a szimulációval kapott logaritmikus hozam időfüggésre. Az ábra így ránézésre viszonylag jó hasonlóságot mutat egy valódi hozam



3.10. ábra. Hozam időfüggése a „csordaszellemű” mikroszkopikus modell eredményeként.

időfüggvénnyel (lásd a 2.1/b) ábrát). Látszik, hogy az extrém ugrások sokkal gyakoribbak, mint egy véletlen bolyongás esetén, azaz a hozam eloszlás a Gaussnál lassabb lecsengést mutat. A másik szembevetendő tulajdonság a valós adatoknál szintén megfigyelt „volatilitás fürtöződés”, tehát az a jelenség, hogy a hozam időfüggésében a nagy ugrások nem rendezetlenül helyezkednek el, hanem egymáshoz közel, fürtöket alkotva, ami a hozam abszolút értéke esetén lassú lecsengésű autokorrelációs függvényt eredményez.

A fenti viselkedéseket, azaz a hozam eloszlást és abszolút értékének auto-korrelációját, a 3.11 ábra mutatja. A legmeglepőbb eredmény, hogy a logaritmikus hozam eloszlására, 3.11/a) ábra, nem egy egyszerű hatványfüggvény lecsengésű eloszlást kapunk, hanem egy eloszlást két különböző exponens értékkel, ami hasonló a valós piacokon megfigyelt jelenséghez [Gopikrishnan et al., 1999]. Ez azonban csak minőségi egyezés, az exponensek értékei nem egyeznek meg a valós exponens értékekkel. A mi esetünkben a kisebb exponens értéke  $\alpha_1 = 1.72$ , ellentétben a valóságban mért 2.4-es értékkel. A nagyobb exponens értéke  $\alpha_2 = 3.36$ , ami a valós esetben 4.



3.11. ábra. *a)* A logaritmusos hozam eloszlása. Log-log skálán ábrázolva két, közelítőleg lineáris tartománya van. Az elsőhöz tartozó exponens értéke  $\simeq 1.72$  míg a másodiké  $\simeq 3.36$  *b)* Az abszolút logaritmusos hozamok összegének szórása,  $D(n)$ . A szaggatott vonal a korrelálatlan,  $\delta = 0.5$ , idősortnak felel meg. A folytonos vonal egy illesztés,  $\delta = 0.95$  exponenssel.

Ahhoz, hogy megvizsgáljuk a hozam abszolút értékének autokorrelációját, tekintjük az abszolút hozamértékek összegének szórásnégyzetét:

$$D^2\left(\sum_{i=1}^n x_i\right) \equiv D^2(n) = n + 2 \sum_{m=1}^{n-1} (n-m) C(m) \sim n^{2\delta}, \quad (3.15)$$

ahol  $C(m) \sim m^{-\kappa}$  az autokorrelációs függvény. A levezetést a B függelék tartalmazza.

$\kappa \geq 1$  esetén  $\delta = 0.5$ , ugyanakkor lassú lecsengésű korreláció,  $\kappa < 1$  esetén  $0.5 < \delta < 1$ , és a korreláció exponense,  $\kappa$  kiszámolható a  $\kappa = 2 - 2\delta$  összefüggésből. A 3.11/b) ábra a varianciát, mint az összegzett abszolút hozamértékek számának függvényét mutatja,  $D(n)$ -et. Nyilvánvalóan látszik, hogy a logaritmusos hozam abszolút értékének időfejlődése korrelált, és autokorrelációja hatványfüggvény lecsengésű, de az exponens értéke,  $\kappa \simeq 0.05$ , ebben az esetben sem egyezik a való adatokon mért értékkel,  $\kappa \simeq 0.3$ .

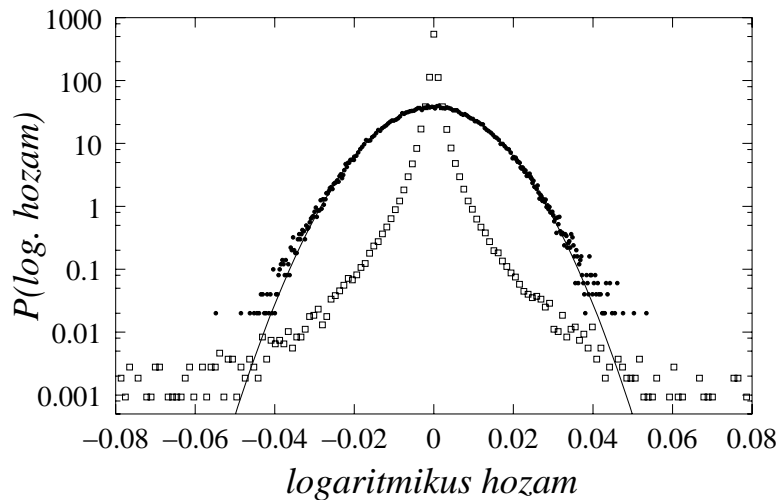
Összefoglalásként a modellünk előnye az egyszerűségében rejlik, és ugyanakkor a valós adatokon megfigyelt „stílusjegyekkel” minőségi egyezést mutatnak. A hátránya, hogy szerepelnek benne paraméterek (például a csoport kialakítási paramétere,  $p$ , az aktivitás paramétere,  $b^*$ ,  $a_m$ ), amelyeket be kell állítani, és az eredmények, azaz az eloszlás és a korreláció exponensei függenek ettől. További vizsgálatok szükségesek ezen paraméterek szerepének tisztázására.

### 3.2.5. Átmenet a Gauss eloszlásba a „csordaszellemű” modellek esetén

A 2.3.2 fejezetben azt láttuk, hogy a perkolációs mikroszkopikus modell esetén, ha az aktív csoportok száma kicsi, az eloszlás hatványfüggvény lecsengésű lesz. A kü-

lönböző, de kis aktivitáshoz tartozó eloszlások egymásra skálázhatóak. Ugyanakkor nagy aktivitás esetén az eloszlás a Gausshoz konvergál, lásd a 2.15 ábrát.

Az általunk használt modellben, [Kullmann and Kertész, 2001c], ahol a csoportok kialakítása nem perkolációval, hanem egy preferált növekedési modellel történik, szintén fellép ez a jelenség. A 3.12 ábrán azt látjuk, hogy míg az aktivitás kicsi, az eloszlás lassú lecsengésű lesz, de megnövelve az aktivitás értékét, a hozamértékek normális eloszlást követnek.



3.12. ábra. A logaritmikus hozam eloszlása különböző aktivitás értékek mellett. A négyzetekkel jelölt görbe abban az esetben adódik, ha az aktivitás minimális, tehát csak egyetlen eladó és egyetlen vásárló csoport létezik. A körökkel jelzett görbe 10000-szer nagyobb aktivitás esetén lép fel. A folytonos vonallal jelzett görbe egy illesztett Gauss eloszlás.

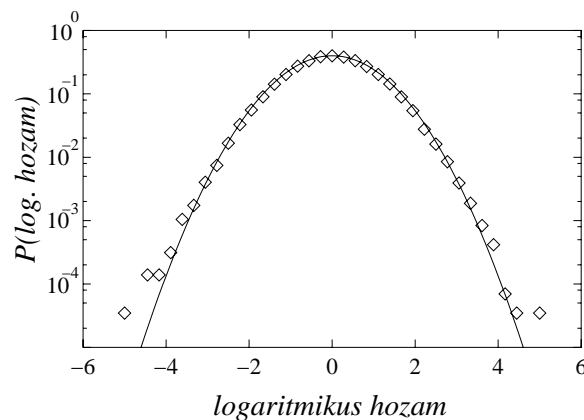
Felvetődik a kérdés, hogy mi okozza a fenti modellekben ezt az átcsapást? Az egyik lehetséges ok az lehet, hogy az eloszlás nem a Lévy tartományban van, azaz véges szórással rendelkezik, így a centrális határeloszlás tétele értelmében a normális eloszláshoz konvergál a valószínűségi változók összegzésével. Ez a jelenség figyelhető meg az általunk alkalmazott modell esetén. A sűrűségfüggvény nagy logaritmikus hozamértékek esetén  $\alpha \simeq 4$  exponenssel cseng le, ami kívül esik a Lévy tartományon, ( $0 < \alpha < 2$ ).

A perkolációs modell esete azonban kevésbé egyértelmű. A perkolációs fürtök eloszlása a Lévy tartományban van, ugyanis  $P(s) \propto s^{-\tau}$ , ahol  $2 < \tau < 2.5$ , így az ilyen típusú független valószínűségi változók összege sokkal inkább egy stabil Lévy eloszláshoz kéne hogy konvergáljon, mint a normális eloszláshoz. Ha a csoportok kialakítása nem perkolációval történik, viszont az eloszlás továbbra is Lévy tartományban van, azt figyelték meg, hogy a Gausshoz való konvergencia nem lép fel [Chowdhury and Stauffer, 1999]. (Itt megjegyezzük, hogy noha a 2.3.1 fejezetben említett perkolációs modellben, [Stauffer and Sornette, 1999] a hozam sűrűségfüggvényének exponensére  $\alpha \simeq 4$  adódott a különböző perkolációs valószínűségekre való

átlagolás miatt, a normális eloszláshoz való konvergenciát abban az esetben is megfigyelték, amikor nem alkalmazták ezt az átlagolást.)

A kérdés az tehát, hogy a megfigyelt normális eloszláshoz való konvergencia a perkolációs fürtök valamely speciális tulajdonságából adódik-e – abból, hogy fürtök nem függetlenek egymástól –, vagy más okra vezethető vissza? Ez a másik ok lehet, hogy annak ellenére, hogy az eloszlás a Lévy tartományban van, létezik a rendszerben egy levágás, ami miatt véges szórása lesz az eloszlásnak, és a centrális határeloszlás tétele érvényesül. A levágás a rendszer véges volta miatt adódhat.

Ahhoz, hogy megvizsgáljuk ezt a kérdést, [Stauffer and Sornette, 1999]-hez hasonlóan szimuláltuk a gazdasági folyamatot, de a csoportokat nem perkolációval, hanem egy Lévy bolyongás segítségével generáltuk. Így ezek mérete egymástól független volt, ugyanakkor továbbra is feltettük, hogy a rendszerben levő résztvevők száma véges, tehát az eloszlásban lesz egy levágás [Kullmann and Kertész, 2001a]. Ez a feltétel a perkolációs modellben is fennállt, hiszen maga a perkolációs rács véges volt. Ellenében a [Chowdhury and Stauffer, 1999] referenciával, mi azt tapasztaltuk, hogy a hozamértékek összegzésének hatására létrejön a normális eloszláshoz való konvergencia, lásd a 3.13 ábrát, ami azt sugallja, hogy az ok a csoportméret eloszlás levágásában keresendő, amely miatt véges szórása lesz az eloszlásnak.



3.13. ábra. Logaritmikus hozam eloszlása abban az esetben, ha a csoportok mérete egy levágással rendelkező Lévy eloszlást követ. A levágás a véges elemszám miatt lép fel. A csoportméret eloszlás exponense  $\alpha = 1.1$ , az aktivitási érték (tehát az egy adott pillanatban aktív csoportok hányada) :  $a = 0.4$ . A levágási értéket  $10^5$ -nek választottuk, ami a legnagyobb lehetséges csoport mérete. A folytonos vonal egy illesztett Gauss eloszlás.

Nyilvánvalóan az eloszlás Gausztól való eltérése, – tehát, hogy mennyi valószínűségi változó összeadása szükséges a Gauss eloszlás eléréséhez –, függ a rendszer méretétől, mivel a levágási érték, azaz a legnagyobb csoport mérete,  $s_m$ , szintén rendszerméret függő. Ugyanakkor felhívjuk a figyelmet, hogy a modellben a konvergencia folyamat kontroll paramétere az aktivitás,  $a$ , és nem az összegzett csoportok száma.

Az aktivitás azon értékének, amelynél a konvergencia fellép, a méretfüggése nem triviális.

Kötött aktivitás érték esetén a hozamérték számításához szükséges összegzendő értékek száma átlagosan lineárisan függ az üzletelő elemek teljes számától, tehát a perkolációs modell esetén a rendszer, vagyis a rács méretétől. Tehát kétszer akkora rendszerméret, de ugyanakkora aktivitási érték esetén a hozam számításánál figyelembe vett elemek száma átlagosan kétszer akkora lesz. Abban az esetben, ha a Gausstól való eltérés,  $\Delta$  szintén lineárisan függ a rendszermérettől, azt kapjuk, hogy az átcsapás (tehát ahol a hozam eloszlások már nem skálázhatóak egymásba, hanem a Gausshoz tartanak), konstans aktivitásnál lép fel, a rendszermérettől függetlenül. A következőkben bebizonyítjuk, hogy a perkoláció esetén ez érvényesül.

A perkolációs elméletben a legnagyobb fűrtméret a rendszermérettől az alábbi módon függ [Stauffer and Aharony, 1994]:

$$s_m \propto L^{d-\beta/\nu} \propto N^{(d-\beta/\nu)/d},$$

ahol  $N$  a rácspontok száma. Jelölje  $p_n(x)$  az  $n$  darab független, azonos eloszlású valószínűségi változók összegének sűrűségfüggvényét. Ennek a normális eloszlástól való eltérése,  $\Delta$ , konstans  $x$  érték esetén felírható az alábbi alakban [Feller, 1968]:

$$p_n(x) - \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \equiv \Delta \propto \frac{1}{n} \frac{\alpha_4}{\alpha_2^2}, \quad (3.16)$$

ahol  $\alpha_4$  a negyedik,  $\alpha_2$  a második centrális momentum, és perkoláció esetén ezek rendszerméret függése a következő lesz:

$$\alpha_4(N) = \sum_{s=1}^{s_m} s^4 n_s \propto s_m^{5-\tau} \propto N^{\frac{(5-\tau)(d-\beta/\nu)}{d}}, \quad (3.17)$$

$$\alpha_2(N) = \sum_{s=1}^{s_m} s^2 n_s \propto s_m^{3-\tau} \propto N^{\frac{(3-\tau)(d-\beta/\nu)}{d}}. \quad (3.18)$$

Ha behelyettesítjük a (3.17) és a (3.18) egyenleteket a (3.16) egyenletbe, megkapjuk az  $n$  valószínűségi változó összegének eloszlása és a Gauss eloszlás közti távolságot, mint a rendszerméret függvényét:

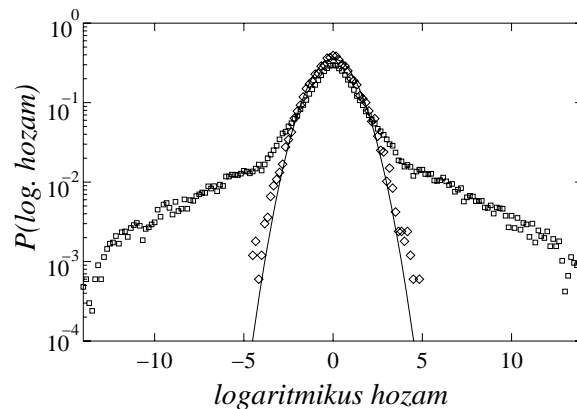
$$\Delta(N) \propto \frac{1}{n} N^{\frac{(d-\beta/\nu)(\tau-1)}{d}} \quad (3.19)$$

Felhasználva a perkolációs skálatörvényeket, [Stauffer and Aharony, 1994], a fenti kifejezés exponensére  $\frac{(d-\beta/\nu)(\tau-1)}{d} = 1$  adódik, ha a dimenzió:  $2 < d < 6$ . Mivel a Cont-Bouchaud modellben az összegzett csoportok átlagos száma arányos a rendszermérettel,  $\langle n \rangle = 2aN$ , így (3.19)-ből adódik:

$$\Delta(N) \propto \frac{1}{a}. \quad (3.20)$$

Ha az átcsapást  $\Delta = \text{const}$  értékkel definiáljuk (tehát az átcsapás ott történik, ahol  $\Delta$  kisebb, mint egy meghatározott konstans érték), (3.20) alapján azt láthatjuk, hogy az átcsapáshoz tartozó aktivitási érték független lesz a rendszer méretétől.

A Gauss eloszláshoz való átcsapásnak egy másik oka is lehet azonban, nevezetesen a kialakított csoport konfigurációk véges száma. A perkolációs mikroszkopikus modellben, [Stauffer and Sornette, 1999], az üzletelési folyamatot egy fix csoportkonfiguráció mellett történt, majd ezután ennek eredményeit átlagolták különböző csoportkonfigurációkra. Abban az esetben, amikor a konfiguráció fix, szintén definiálhatunk egy levágási értéket a méreteloszlásban, ami az adott konfigurációban szereplő legnagyobb csoport mértete lesz. Ez a levágás akkor is fellép, ha végtelen nagy rácson szimulálunk véges számú fürtöt. Ha nem átlagolunk ki különböző mintákra, az eloszlás Gausshoz fog konvergálni, mivel a levágás miatt véges lesz az eloszlás szórása. Ez a jelenség akkor is jelentkezik, ha ugyan kiátlagolunk különböző mintákra, de a minták száma kevés. Minél több mintára történik az átlagolás, a legnagyobb csoport okozta levágás értéke „elkenődik”, és a végeredmény egy hatványfüggvény lecsengésű eloszlás lesz, ahogy azt a 3.14 ábra is szemlélteti. Természetesen abban az esetben, ha egy végtelen nagy rendszert szimulálunk és a konfigurációk száma, amire az átlagolás történik szintén végtelen, az így kialakuló hozam eloszlása perkoláció esetén stabil Lévy eloszlást eredményez.



3.14. ábra. A hozam eloszlás Gauss jellegű viselkedése, a minták korlátozott száma miatt. A fenti példában a csoportméret eloszlás Lévy eloszlású,  $\alpha = 1.1$  exponenssel. Az aktivitás értéke  $a = 0.4$ . A ( $\diamond$ ) 10 mintára történő átlagolás eredménye. Az eloszlásra jól illeszkedik a Gauss eloszlás. A ( $\circ$ ) esetében 100 mintára átlagoltuk a folyamatot, ebben az esetben az eloszlás lassan lecsengő jellegű mutat.

## 4. fejezet

# Részvény árfolyamok keresztkorrelációi

### 4.1. Időfüggetlen keresztkorrelációk vizsgálata

A 2.2 fejezetben említettük, hogy a kockázat csökkentésének egyik hatékony módja a portfólió összeállítása, tehát nem egyetlen részvény, hanem egy részvény csomag vásárlása. Ahhoz, hogy meghatározzuk, milyen részvények kerüljenek a csomagba, ismernünk kell ezek egymáshoz viszonyított időbeli változását. Ennek legegyszerűbb meghatározása a keresztkorrelációk mérésével történik. A keresztkorreláció felfogható, mint a részvények közti effektív kölcsönhatás erőssége, amely alapján a részvények csoportokba rendezhetőek. (Itt a csoportok nem a portfólió csomagot jelölik, csupán arról van szó, hogy úgy kívánjuk rendezni a részvényeket, hogy a „hasonlóak” ugyanazon csoportba kerüljenek.) Minél nagyobb a korreláció két részvény között, annál hasonlóbbnak nevezzük őket, és annál nagyobb valószínűséggel kerülnek ugyanazon csoportba. Nyilvánvaló, hogy a csoportba rendezés nem csak egyféleképpen végezhető el, hanem függ attól, hogy milyen kölcsönhatási erősséget kívánunk meg ahhoz, hogy két elemet ugyanazon csoportbelinek tekintsünk. Ez kapcsolatban van azzal, hogy a portfólió elméletben mekkora kockázat mellett kívánjuk maximalizálni a nyereségünket. A csoportoknak ilyenmódon egy hierarchikus struktúrája áll elő.

Ebben a fejezetben ennek a struktúrának a meghatározásával foglalkozunk. A bevezetőben említettük, hogy Mantegna az MST módszerével oldotta meg a részvénycsoportok közti hierarchikus rend kimutatását. Mi egy másik, szintén a bevezetőben említett módszerrel, a szuper-paramágneses csoportosítással határozzuk meg ezt, amely egy  $q$  állapotú Potts modell vizsgálatára vezethető vissza.

### 4.1.1. Szuper-paramágneses Potts modell alkalmazása

Emlékeztetve a bevezetőben írtakra, a szuper-paramágneses csoportosítási módszer lényege, hogy a (2.32) egyenlet által meghatározott Hamilton függvénnyel jellemzett rendszernek mérjük a szuszeptibilitását,  $\chi(T)$  (lásd a (2.34) egyenletet) a hőmérséklet függvényében, és meghatározzuk a különböző szuper-paramágneses fázisokhoz tartozó hőmérséklet tartományokat. Az így meghatározott hőmérsékleteken (például a hőmérséklettartomány mértani közepe) kiszámolt spin-spin korrelációs függvény,  $\langle \delta_{s_i, s_j} \rangle$ , fogja megadni, hogy a halmaznak mely elemei tartoznak egy csoportba. Az eredeti Domany modellben a halmaz elemeinek egy metrikus térben elhelyezkedő pontok feleltek meg, és az elemek közti kölcsönhatás a pontok közti távolság függvénye volt, ahogy azt a (2.33) egyenlet definiálja.

Egy fontos probléma felmerült a mi esetünkben. A részvények közt nem csak „vonzás” (pozitív korreláció) lehetséges, hanem „taszítás” (negatív vagy antikorrreláció) is. Sőt, bizonyos szempontból ez a „taszító” hatás fontosabb is lehet a portfólió elmélet szempontjából, hiszen a kockázat minimalizálásának egyik legegyszerűbb módszere, ha egymással ellentétesen mozgó részvényeket válogatunk össze, ami redukálja a nagy árzuhanás esetén kialakuló veszteséget. Az antikorrreláció kezelésének a problémáját úgy oldottuk meg, hogy elvetettük a metrika feltételét, ezáltal lehetőség nyílt, hogy negatív (anti-ferromágneses) csatolásokat is megengedjünk. A Potts spinek közti csatolás a mi esetünkben explicit függvénye a részvényárfolyamok közti korrelációs együtthatónak,  $\rho_{ij}$ :

$$J_{ij} = \text{sgn}(\rho_{ij}) \left( 1 - \exp \left\{ -\frac{n-1}{n} \left[ \frac{\rho_{ij}}{a} \right]^n \right\} \right) \quad (4.1)$$

A korrelációs együtthatók meghatározását a (2.30) egyenlet definiálja. Fontos, hogy a kölcsönhatás gyorsan növekvő függvénye legyen a korrelációknak. Ezzel ugyanis a probléma hosszú-hatótávolságú jellegét (hiszen minden részvény minden másikkal kölcsönhat, még ha gyengén is) rövid hatótávolságúvá teszi, azaz a gyenge korrelációkat még gyengébb, szinte nulla erősségű kölcsönhatással helyettesíti. Ez nem csak a számítási igény lecsökkentéséhez szükséges, hanem ahhoz is, hogy kialakuljon a szuper-paramágneses fázis. Ahhoz ugyanis, hogy a különböző spin csoportok egymástól függetlenek legyenek (egymástól függetlenül meg lehessen őket forgatni), szükséges, hogy a csoportok közti kölcsönhatás nagyon kicsi, vagy nulla legyen. Ennek megértéséhez térjünk vissza az eredeti problémához, amikor a halmaz elemei közti hasonlóságot mint a távolság függvényét tekintjük. Tegyük föl, hogy három pontunk van, ezek közül kettő,  $A$  és  $B$  közel egymáshoz, míg a harmadik,  $C$ , távolabb tőlük. Akkor fog a két közeli pont,  $A$  és  $B$ , egy csoportba tartozni és a harmadik,  $C$ , egy másik csoportba, ha egy adott hőmérsékleten annak a valószínűsége, hogy  $C$  másik irányba mutat mint  $A$  vagy  $B$ :  $P_{AC} \sim \exp\{-J_{AC}/T\}$ , nagy, míg annak a valószínűsége, hogy  $A$  és  $B$  ellenkező irányba mutatnak:  $P_{AB} \sim \exp\{-J_{AB}/T\}$  kicsi. Minél

gyorsabban cseng le a távolsággal a csatolási erősség, annál könnyebben alakul ki a  $P_{AB} \ll P_{AC}$  feltétel.

Egy másik ok, ami miatt szükséges, hogy a csatolási együttható gyorsan lecsengő függvénye legyen a korrelációknak az, hogy a kis korrelációk inkább a zaj mint a valós információ következményei, és nem akarjuk, hogy a zaj befolyásolja az eredményeinket. A (4.1) egyenletben definiált csatolási állandónak két paramétere van,  $a$  és  $n$ . Ezeket úgy kell megválasztani, hogy létezzen a szuper-paramágneses fázis, ugyanakkor az eredményeink ezen belül ne függjenek túl erősen a paraméterek választásától. A paraméterek finomhangolása arra is szolgál, hogy könnyebben megfigyelhető legyen az átalakulási pont, azaz a szuszceptibilitás függvényben élesebb csúcsok legyenek és a köztük lévő lapos tartomány nagyobb legyen. Az  $a$  paraméter egy lehetséges választása az egy spinhez tartozó legnagyobb korrelációs értékek átlaga:  $a = 1/N \sum_{i=1}^N \max_j(\rho_{ij})$ . A (4.1) egyenletben szereplő  $n$  hatvány a kölcsönhatás tartományát állítja be, az exponensben szereplő  $n/(n-1)$ -es faktor a kölcsönhatási függvény inflexió pontját tolja el.

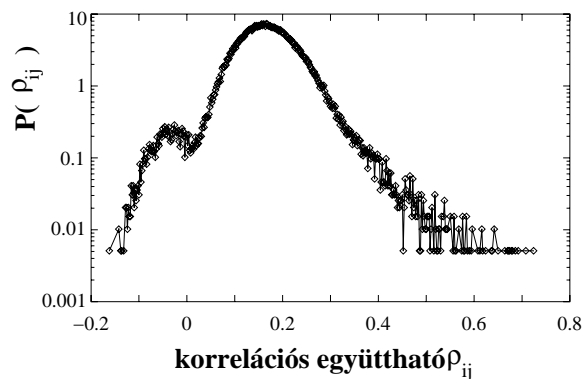
A rendparamétert a (2.35) egyenlet definiálja, ahol  $N$  a spinek (mi esetünkben vállalatok) száma,  $q$  a lehetséges állapotok száma, amit egy spin felvehet. Illetve  $N_{max}$  a maximális száma azon spineknek, amelyek egy irányba mutatnak. A  $q$  paraméter beállítására nincs előírt recept, mert mindig az adott problémától függ az értéke. Túl kis  $q$  érték megnehezíti az SPM csoportok meghatározását mivel ekkor a különböző csoportok arra kényszerülnek, hogy ugyanabba az irányba mutassanak. Ugyanakkor túl nagy  $q$  érték esetén a szimulációs idő nő meg, ami az egyensúly beállításához kell. Az eredmény azonban nem függ túlságosan a  $q$  értékétől. Egy kézenfekvő mód a  $q$  paraméter meghatározására, ha egyenlővé tesszük az elemek számával:  $q = N$ .

A paraméterek meghatározása után a rendszert szimulálni kell a hőmérséklet függvényében. Az elsőrendű SPM fázisátalakulások azon hőmérsékletnél lesznek, ahol a szuszceptibilitásban  $\chi(T) = N(\langle m^2 \rangle - \langle m \rangle^2)/T$  egy csúcs található. Kényelmi szempontból mi a  $\hat{\chi}(T) = T\chi(T)/N$  függvényt használtuk. Ennek az az előnye, hogy ez a függvény a két fázisátalakulási csúcs közt majdnem konstans a hőmérséklet függvényében, ezáltal még inkább megkönnyíti a különböző SPM fázisok detektálását. A legegyszerűbb, nem triviális esetben, tehát amikor létrejön egy SPM fázis, összesen két fázisátalakulás történik: a Ferromágneses (FM)  $\implies$  szuper-paramágneses (SPM) fázisok, és a szuper-paramágneses (SPM)  $\implies$  Paramágneses (PM) fázisok közt. De lehetséges hogy az SPM fázison belül is találunk újabb csúcsokat a szuszceptibilitásban, amelyek azt jelentik, hogy több mint egy karakterisztikus rendeződés alakul ki. A hőmérséklet növelésével a különböző csoportok alcsoportokra esnek szét, azaz a csoportok egy hierarchikus struktúrája áll elő. A vizsgálódás célja ezen hierarchikus struktúra kimutatása.

## Felhasznált adatok

Két rendszert vizsgáltunk. Az egyik a Dow Jones Industrial Average (DJIA) tőzsdeindexből kiválasztott 30 vállalat, a másik a Standard & Poor's 500 (S&P500) index vállalatai közül választott 443 vállalat volt. Az vállalatok hozamai közti korrelációt az 1989 július 3 és 1995 október 27 közti időintervallumra átlagoltuk. Ez az átlag adja meg a korrelációs együtthatókat,  $\rho_{ij}$ , amelyekből a csatolási állandót számoljuk.

Míg a DJ vállalatok közt az adott időtartományban csak pozitív korreláció létezik,  $\rho_{ij} \geq 0 \forall i, j$  esetén, (azaz a spinek közt csak FM-es kölcsönhatás alakul ki) az S&P500 vállalatai közt létezik negatív korreláció is, amely lehetőséget nyújtott, hogy megvizsgáljuk, hogy az AFM-es kölcsönhatás milyen mértékben befolyásolja a kialakult csoport struktúrát. A 4.1 ábra a keresztkorrelációk együtthatóinak eloszlását ábrázolja. Az ábráról azt olvashatjuk le, hogy bár léteznek negatív korrelációk, de ezek száma és nagysága sokkal kisebb volt, mint a pozitív korrelációs együtthatóké.



4.1. ábra. S&P500 tőzsdeindexből kiválasztott 443 vállalat közt mért korreláció együtthatóinak eloszlása. A korrelációk körülbelül 6 év időintervallumra lettek kiátlagolva. Az ábra azt mutatja, hogy léteznek ugyan negatív korrelációs együtthatók, de ezek nagysága sokkal kisebb, mint a pozitív együtthatóké, valamint számuk is sokkal kevesebb.

### 4.1.2. Eredmények értékelése

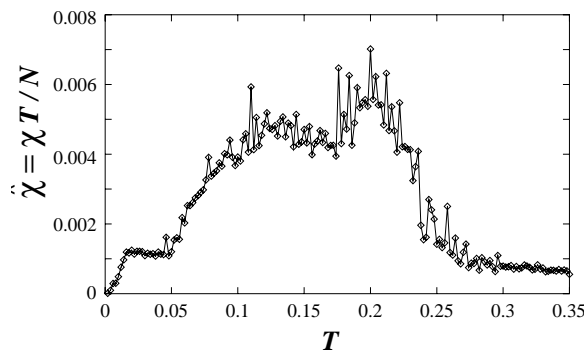
Elsőként tehát a DJ indexet vizsgáltuk meg, amely  $N = 30$  vállalatot tartalmaz. Említettük, hogy ennél a halmaznál csak FM-es kölcsönhatás létezik, ezért a szimuláció a hatékony Swendsen-Wang algoritmussal [Swendsen and Wang, 1987] végezhető. A (4.1) kölcsönhatás paramétereit  $a = 0.43$  illetve  $n = 8$  értékeknek választottuk. A szuszceptibilitás  $\hat{\chi}$  hőmérséklet függését a 4.2 ábra szemlélteti. Mivel a rendszer mérete kicsi – összesen 30 elemet tartalmaz –, nagy fluktuációk lépnek fel a szuszceptibilitás hőmérsékletfüggésében, ami megnehezíti a csoportok meghatározását. Mint ahogy a (4.2) ábra mutatja, nem teljesen egyértelmű mely csúcsok jelzik a fázisátalakulást, és melyek azok, amelyek csak a fluktuáció következményei. Ugyanakkor,

mint említettük, két fázisátalakulási hőmérséklet közt a  $\hat{\chi}$  szuszeptibilitás majdnem konstans értékű. Ezt a tulajdonságot használtuk ki a különböző fázisok meghatározásához, amely alapján négy különböző hőmérsékleten vizsgáltuk meg a spin-spin korrelációs függvényt. Ennek eredményét a 4.1 táblázat mutatja. (A rövidítések a vállalatok szokásos szimbólumai, amelyeket a C függelékben találjuk meg.) A táblázatban csak azon csoportokat tüntettük fel, amelyeknek háromnál több elemük volt. A 4.1 táblázatból azt olvashatjuk le, hogy a DJ vállalatai esetében összesen két cso-

4.1. táblázat. A DJ vállalatai közt kialakuló csoportok a különböző felbontás, azaz hőmérséklet mellett.

$T=0.003$	(1)	AA, ALD, AXP, BA, BS, CAT, CHV, DD, DIS, GE, GM, IP, JPM, KO, MCD, MMM, MO, MRK, PG, T, TX, UK, UTX, XON
$T=0.13$	(1)	AA, DD, GE, GM, IP, JPM, KO, MCD, MMM, MO, MRK, PG, T
	(2)	CHV, TX, XON
$T=0.15$	(1)	CHV, TX, XON
	(2)	DD, GE, IP, KO, MCD, MMM, MO, MRK, PG, T
$T=0.28$	(1)	CHV, TX, XON

portot különböztethetünk meg: egy kisebbet, amelyben az olaj-ipari cégek tartoznak, és egy nagyobbat. Az utóbbi a hőmérséklet növekedtével (tehát a felbontás növelésével) kisebb csoportokra hullik szét. A hierarchia ugyan létezik valamilyen mértékben, hiszen  $T = 0.03$  hőmérsékleten még mindkét csoport egybe tartozott, de nem jelentős. Összehasonlítva az eredményeinket a Mantegna által vizsgált eredményekkel [Mantegna, 1999] nagymértékű egyezést találunk.

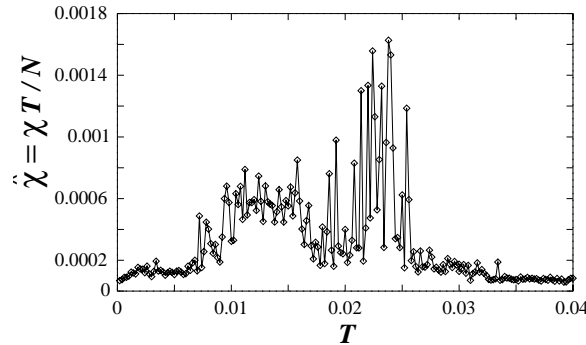


4.2. ábra. A szuszeptibilitás  $\hat{\chi} = T\chi/N$  hőmérséklet függése a DJ tőzsdeindex vállalatai esetében. A szimuláció paramétereit:  $a = 0.43$ ,  $n = 8$  és  $q = 100$  voltak.

A másik általunk vizsgált rendszer a Standard & Poor's 500 tőzsdeindex volt, amelynek tanulmányozása érdekesebbnek mutatkozik, egyrészt a vállalatok nagyobb száma miatt,  $N = 443$ , másrészt ezen vállalatok korrelációi közt már találhatunk negatív értékeket is.

Ahhoz, hogy a „taszítás” hatását megvizsgálhassuk, két különböző módon végeztük a számításainkat. Először csak tiszta ferromágneses kölcsönhatást vettünk figyelembe (a korrelációs együtthatók abszolút értékét tekintettük), másodszer viszont figyelembe vettük a korrelációs együttható előjelét is. Azt találtuk, hogy a két eset közti különbség létezik, de a mi adathalmazunk esetében csak az alapállapotban,  $T = 0$  figyelhető meg.

A FM esetben a szuszeptibilitás  $\hat{\chi}$  hőmérséklet függését a (4.3) ábra szemlélteti. A szuszeptibilitásban itt is nagy fluktuációk lépnek fel, amelyek abból adódnak, hogy a 4.1 ábrán látható korrelációs együtthatók eloszlása keskeny. Emiatt a csoportok szétválasztása érzékeny lesz a hőmérsékletre.

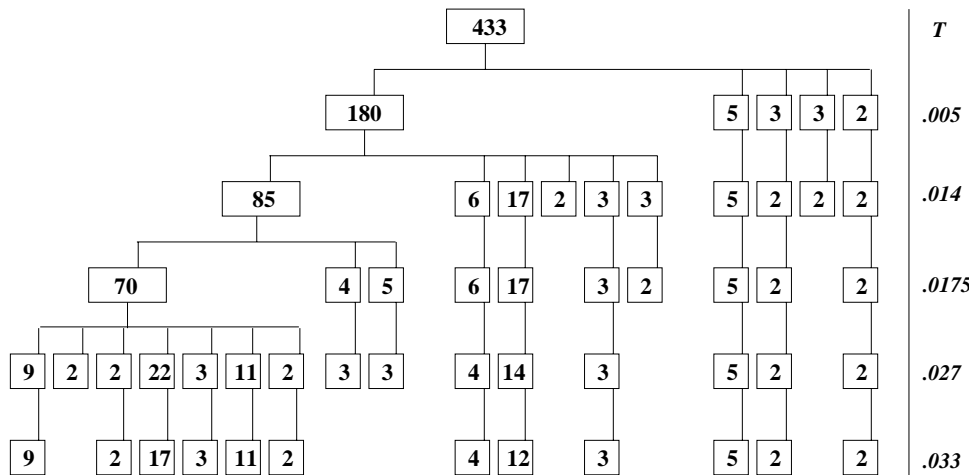


4.3. ábra. A szuszeptibilitás  $\hat{\chi} = T\chi/N$  hőmérséklet függése az S&P500 tőzsdeindex vállalatai esetében. A szimuláció paramétereit:  $a = 0.65$ ,  $n = 10$  és  $q = 50$  voltak.

Az S&P500 vállalatok alkotta rendszer esetében világosan kirajzolódik a csoportok közt kialakuló hierarchikus rend, ezt mutatja a 4.4 ábra. A téglalapokban szereplő számok azt mutatják, hogy hány vállalat tartozik egy csoportba. A módszer hatékonyságát az mutatja, hogy az eredményeink jól egyeznek a MST módszerrel kapott eredményekkel [Mantegna, 1999], és nem kizárólag a hierarchikus struktúrát illetően, hanem az egyes csoportok tartalmát tekintve is. Ez annyit jelent, hogy a hierarchia legalsó szintjén (a legmagasabb hőmérsékletet) a különböző csoportok a gazdaság egy-egy ágazatának felelnek meg. Az olaj-, elektromos áram-, arany bányászati vállalatok különböző csoportokat alkotnak.

A fenti FM képből félrevezethető eredményt adhat, hogy nem tettünk különbséget a vállalatok közti „vonzás” és „taszítás” közt, azaz erős antikorrrelációkat és erős korrelációkat egyenértékűnek vettünk. A portfólió optimalizáció szempontjából a „taszításnak” különösen jelentős szerepe van. Tehát meg kell vizsgálnunk a fenti rendszert abban az esetben, amikor a korrelációk előjelét is figyelembe vesszük.

Egy inhomogén, hosszú-hatótávolságú Potts modell szimulációja, ahol a kölcsönhatások váltakozó előjelűek, nem egyszerű, még a rendparaméter definíciója sem triviális. A mi esetünkben szerencsére a negatív kölcsönhatások másodrendű szerepet játszanak összehasonlítva a FM-es kölcsönhatásokkal: gyengébbek és kevesebb da-



4.4. ábra. Az S&P500 tőzsdeindex vállalatai esetében a csoportok közt kialakuló hierarchikus struktúra szemléltetése. A téglalapokban az ugyanazon csoportba tartozó vállalatok száma szerepel. Ahogy nő a hőmérséklet úgy esnek szét a csoportok kisebb és kisebb alcsoportokra. (Azokat a csoportokat, amelyek csak egy vállalatot tartalmaznak, nem tüntettük fel.)

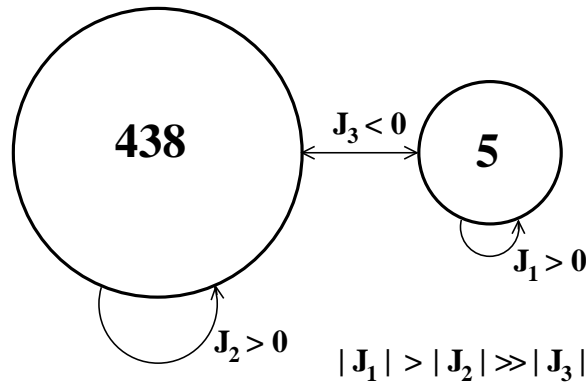
rab van belőlük. Erre utal a 4.1 ábra, amely a korrelációs együtthatók eloszlását szemlélteti. További egyszerűsítést jelent, hogy szinte az összes negatív kölcsönhatás mindössze 5 vállalathoz tartozik – az arany bányászattal foglalkozó vállalatokhoz –, amelyek ugyanakkor nagyon erős kölcsönhatásban vannak egymással. A FM-es esetben még a legnagyobb hőmérsékleten is egy csoportot alkotnak. Azt várjuk tehát, hogy a nagyon alacsony hőmérsékletek kivételével nem lesz szignifikáns különbség a FM-es és az AFM-es eset közt. Ennek szellemében a szimuláció egy egyszerű Metropolis algoritmussal is elvégezhető, nincs szükség szofisztikáltabb eszközökre, mint például a multi-kanonikus algoritmus [Berg and Neuhaus, 1992]. Ennek ellenére az alacsony hőmérsékletű konfiguráció meghatározása sem egyértelmű. A rendszer könnyen egy lokális minimumba kerülhet és nagyon sok szimulációs idő szükséges, hogy újra lépjen onnan. Ezért egy szimulált hőkezeléshez [Kertész and Kondor, 1996] hasonló módszert alkalmaztunk. A rendszert egy magasabb hőmérsékletre gerjesztjük, majd a hőmérsékletet fokozatosan csökkentjük úgy, hogy minden hőmérsékleti szinten azon konfigurációt tartjuk meg, amelynek az energiája a legkisebb volt.

4.2. táblázat. Energia értékek és a hozzájuk tartozó konfigurációk alacsony hőmérsékleten, az S&P500 index vállalatainak vizsgálata esetében, AFM kölcsönhatások figyelembe vételével.

konfiguráció	vállalatok száma	energia
$C_1$	438, 5	-672,37417
$C_2$	443	-672,35707

A 4.2 táblázat a konfigurációkat és az energia értéküket tartalmazza, amelyet ezen

az alacsony hőmérsékleten kaptunk. A FM-es esetben a legalacsonyabb hőmérsékleten kapott konfiguráció megfelel a táblázatban szereplő  $C_2$  konfigurációnak, amely azonban itt nem a legalacsonyabb energiával rendelkezik. Tehát az AFM eset alapállapota különbözik a FM eset alapállapotától, az előbbi esetben nem az összes vállalat tartozik egy csoportba. Ez azonban nem meglepő, hiszen ezt az eredményt akár szimuláció segítsége nélkül is ki lehet következtetni, pusztán a korrelációs tulajdonságok alapján. Erre szolgál a 4.5 ábrán vázolt sematikus rajz.



4.5. ábra. Sematikus ábra, amely szemlélteti, hogy az S&P500 vállalatok esetén, AFM esetben, alacsony hőmérsékleten milyen lesz a konfiguráció, és ez hogyan állapítható meg pusztán a kölcsönhatások figyelembevételével. Az ábra jobb oldalán szereplő körben az aranybányászattal foglalkozó vállalatok szerepelnek. Ezek közti kölcsönhatás a legerősebb. Ezek és a többi vállalat közt majdnem az összes kölcsönhatás negatív.

A célunk egyrészt az volt, hogy meghatározzuk a vállaltok közt kialakuló csoportokat két rendszer, a DJ és az S&P500 indexek esetében a szuper-paramágneses csoportosítás [Blatt et al., 1996] eljárásának alkalmazásával. Ez az eljárás ugyanis alkalmas arra, hogy megkülönböztessük a pozitív illetve a negatív korrelációkat. Abban az esetben, amikor csak pozitív korreláció van – DJ vállalatok, illetve az S&P500 vállalatai FM esetben – azt tapasztaljuk, hogy az általunk kapott csoportok jól meggyeznek a korábbi eredményekkel [Mantegna, 1999]. Másrészt megvizsgáltuk, hogy figyelembe véve a negatív kölcsönhatásokat is, hogyan módosul a kialakult struktúra. Azt tapasztaltuk, hogy van eltérés a két eset – FM és AFM esetek – közt, igaz csak gyenge, ami a kis számú és gyenge negatív korrelációknak köszönhető.

## 4.2. Időfüggő keresztkorrelációk vizsgálata

Az előző fejezetben a különböző részvények közti egyidejű korrelációkat vizsgáltuk, amelyeket mint a részvények közti kölcsönhatás mértékét alkalmaztuk. A portfólió optimalizáció szempontjából azonban nem csak az a fontos, hogy megtudjuk, mely részvények állnak kölcsönhatásban egymással, hanem hogy megvizsgáljuk ennek a

kölcsönhatásnak az irányát is. Egy olyan módszert keresünk, amellyel kvantitatív módon jellemezni lehet a kölcsönhatás okát. Ezt ugyanis egyrészt okozhatja egy külső effektus, ami egyaránt és egyszerre hat a két kölcsönható részvényre. Másrészt azonban az is előfordulhat, hogy az egyik vállalat befolyásolja a másikat, és ez a befolyás a másik részvény árának változásában mutatkozik meg. Az egyik legalkalmasabb módszer a kölcsönhatás leírására az időfüggő korrelációk szimmetriájának vizsgálata.

A statisztikus fizikában az időfüggő korreláció fontos szerepet játszik. A Fluktuáció Disszipáció Tételben betöltött szerepe miatt az egyik legfontosabb mennyiség amely meghatározza a transzport együtthatókat. A híres Onsager relációk alapja az időfüggő keresztkorrelációk szimmetria-tulajdonságaiban rejlik. Nyilvánvalóan egy gazdasági rendszerben nem kell megkövetelnünk, hogy az időtükrözési szimmetria vagy a részletes egyensúly teljesüljön. Amint azt a továbbiakban látni fogjuk, a részvények közti korrelációk esetében sok esetben a korreláció maximuma nem az origóban van, hanem attól kissé eltolva, amely azt jelenti, hogy létezik a vállalatok közt egy húzó erő, azaz az egyik befolyással van a másik árváltozására [Kullmann et al., 2002]. Azt is látni fogjuk, hogy ez a húzó effektus nem túl erős, és a korrelációs függvény origótól való eltolódása sem nagy, amelyek oka elsősorban a piac hatékonyságában rejlik, ugyanis ellenkező esetben ez az effektus arbitrage üzletekhez vezethetne.

#### 4.2.1. Felhasznált adatok

A hatékony piacok egyik alaptulajdonsága, hogy a részvény logaritmikus hozamának autokorrelációi exponenciálisan csengenek le, nagyon rövid, néhány perces karakterisztikus idővel. Ebből következően, valamint a bevezetőben ismertetett gondolatmenet alapján azt várjuk, hogy az általunk keresett effektus (korrelációs maximum eltolódás) sem lehet sokkal nagyobb ennél a néhány perces skálánál. Ezen megfontolás miatt szükséges, hogy a vizsgálatot nagy frekvenciás adatokon végezzük. Ez alatt olyan adathalmazt értünk, amelyben a részvény árváltozásai közt eltelt idő néhány másodperc nagyságrendű. A Trade and Quote (TAQ) adatbázist használtuk fel, amely a New York-i tőzsde részvényeinek adatbázisa. Egy 54 napos időintervallumot vizsgáltunk, 1997 december 1 és 1998 március 9-e közt, amely nagy frekvenciás adatokat tartalmazott körülbelül 10000 vállalatra. Mivel ez az időintervallum meglehetősen rövid, csak azon vállalatokat vizsgáltuk, amelyekkel 15000-nél több alkalommal üzleteltek, így a vállalataink száma 195-re redukálódott.

Az üzletkötések egy részvény esetén nem azonos időnként történnek, így ha egy adott időintervallumhoz tartozó hozamot (vagy a logaritmikus hozamot) akarunk definiálni, valahogyan definiálnunk kell az árat két üzletkötés közt is. A tőzsde szabályaihoz igazodva ezt konstansnak vettük, tehát feltételeztük, hogy az ár addig nem változik, amíg nem történik új üzletkötés. A teljes  $T$  időintervallumot  $n$  kis intervallumra osztottuk, ami által egy intervallum mérete nyilvánvalóan  $\Delta t = T/n$ . Ha a

$t$ -edik időpillanatban történik üzletkötés, akkor a logaritmikus hozam értéke:

$$r_{\Delta t}(t) = \ln \left[ \frac{p(t)}{p(t - \Delta t)} \right],$$

más esetben nulla.

A különböző napok közt fellépő hozamértékeket – azokat, amelyek az egyik nap nyitó és az előző nap záró ára közt lépnek fel – figyelmen kívül hagytuk, mert ezek a nagy hozam értékek meghamisíthatnák az eredményeket. Tehát a különböző napokat egymástól függetlennek tekintettük, emiatt a korrelációban az átlagolás két részből áll: először az egy napon belüli időintervallumra, ezután a különböző napokra végeztük el.

### 4.2.2. Zajredukció

A célunk, hogy megvizsgáljuk a különböző részvények közti kereszt-korrelációt, mint a részvények árfolyamai közti időeltolás,  $\tau$  függvényét. Az időfüggő keresztkorreláció,  $C_{A,B}(\tau)$ , a definíció alapján:

$$C_{\Delta t}^{A,B}(\tau) = \frac{\langle r_{\Delta t}^A(t) r_{\Delta t}^B(t + \tau) \rangle - \langle r_{\Delta t}^A(t) \rangle \langle r_{\Delta t}^B(t + \tau) \rangle}{\sigma_A \sigma_B}, \quad (4.2)$$

ahol  $\sigma^2 = \langle (r_{\Delta t}(t) - \langle r_{\Delta t}(t) \rangle)^2 \rangle$  a logaritmikus hozam szórásnégyzetét jelöli. A  $\langle \cdot \rangle$  jelölés a már említett átlagolás a  $T$  időintervallumra.

Felmerül a kérdés, hogy milyen  $\Delta t$  időintervallumot válasszunk. Mivel két üzletkötés közti legkisebb időintervallum egy másodperc, a  $\Delta t = 1s$  választás kézenfekvő megoldásnak tűnne. Ugyanakkor egy ilyen kis ablak esetén elég gyakran előfordulhat, hogy az adott időlépés közben nem történik tranzakció valamelyik részvény (vagy mindkettő) esetén, így a hozamnak a korrelációs függvényhez való hozzájárulása nulla lesz. Mivel a nem nulla hozzájárulások száma kicsi, a korrelációs együttható erősen fluktuálni fog az időeltolás  $\tau$  függvényében. Hogy elkerüljük ezt a problémát, meg kell növelni az időintervallum,  $\Delta t$  méretét, és még tovább javíthatjuk a helyzetet, ha kiátlagolunk a hozamok kezdő időpontjaira. Ekkor a (4.2) egyenletben szereplő átlagolás a következőképpen szemléltethető:

$$\langle r_{\Delta t}^A(t) r_{\Delta t}^B(t + \tau) \rangle = \frac{1}{T} \sum_{t_0=0}^{\Delta t-1} \sum_{k=1}^{T/\Delta t} r_{\Delta t}^A(t_0 + k\Delta t) r_{\Delta t}^B(t_0 + k\Delta t + \tau), \quad (4.3)$$

ahol az első összegzés a hozam kezdő pontjaira történik a másik pedig a teljes időtartomány  $\Delta t$  széles ablakain fut végig.

### Perzisztens véletlen bolyongás

A következőkben vizsgáljuk meg a fenti effektus helytállóságát, azaz azt, hogy növelve a hozam időintervallumát könnyebbé válik a korrelációs csúcsok detektálása, – azaz a korrelációs függvény maximum helyének a lokalizálása. Két mesterséges adatsort szimuláltunk. Az egyik egy egydimenziós Perzisztens Véletlen Bolyongás (PRW) [Furth, 1917, Weiss, 1994], amely annyiban tér el a hagyományos véletlen bolyongástól, hogy annak a valószínűsége,  $\alpha$ , hogy a megelőzővel azonos irányba ugrik, nagyobb, mint az ellenkező irányba történő ugrás, azaz  $\alpha > 0.5$ . Ez annyit jelent, hogy a bolyongó részecske emlékszik a múltjára. A  $t$  időpillanatban az  $x(t) \in \pm 1$  ugrás valószínűsége:

$$P(x(t)) = \alpha \delta_{x(t), x(t-1)} + (1 - \alpha)(1 - \delta_{x(t), x(t-1)}). \quad (4.4)$$

A másik mesterséges idősort az elsőből egyszerűen úgy generáltuk, hogy eltoltuk az időben  $\tau_0$  értékkel és hozzáadtunk egy véletlen zajt, amely normális eloszlású, nulla várhatóértékkel és  $\sigma$  szórással:

$$y(t) = x(t - \tau_0) + \xi(t), \quad \xi \in N(0, \sigma) \quad (4.5)$$

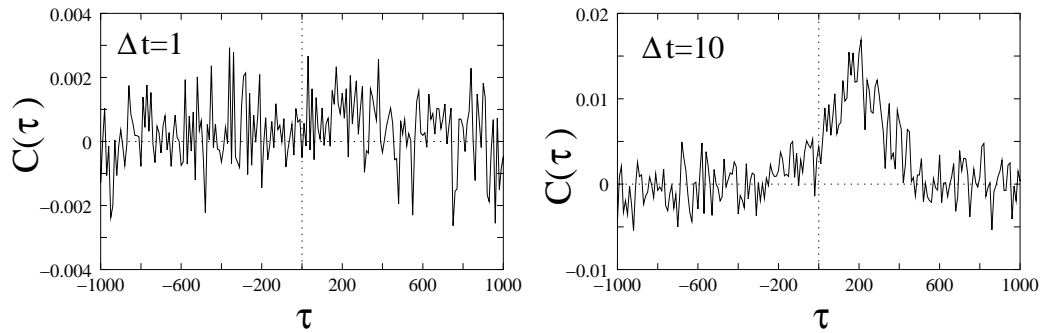
A modell előnye, hogy a korrelációs függvény analitikusan kiszámítható, és ebből meghatározható a korreláció maximális értéke  $\tau_0$ -ban:

$$C(\tau) = \frac{(2\alpha - 1)^{|\tau - \tau_0|}}{\sqrt{\sigma + 1}} \quad (4.6)$$

Miután legeneráltuk a két idősort, véletlenszerűen eldobtunk pontokat mindkét idősorból (a pontok  $\rho$  hányadát tartottuk meg), hogy ugyanaz a probléma merüljön fel, mint a valós adatok esetében, tehát, hogy az ugrások nem egy időben történnek. A 4.6 ábráról nyilvánvalóan látszik, hogy növelve a  $\Delta t$  időkülönbséget, könnyebb meghatározni a korreláció maximum helyét. (Az a tény, hogy eldobtunk néhány pontot kicsit megváltoztatta a maximum értékét a (4.6) egyenlethez képest.)

A 4.6 ábra azt mutatja, hogy a korreláció lecsengése ellentétben a (4.6) egyenlettel nem exponenciálisan, hanem közelítőleg lineárisan cseng le. Ez az átlagolási eljárás miatt van, amit a megnövelt  $\Delta t$  időkülönbség esetén használtunk. Egy nagyobb időkülönbséghez tartozó korrelációs függvény,  $C_{\Delta t}(\tau)$ , felírható mint az egységnyi időkülönbséghez,  $\Delta t = 1$  tartozó korrelációs függvényeknek,  $C_1$ , súlyozott átlaga:

$$\begin{aligned} \langle r_{\Delta t}^A(t) r_{\Delta t}^B(t + \tau) \rangle &= \left\langle \sum_{i=1}^{\Delta t} \delta^A(t + i) \sum_{i=1}^{\Delta t} \delta^B(t + i + \tau) \right\rangle = \\ &= 1C_1(\tau - \Delta t + 1) + \dots + (\Delta t - 1) C_1(\tau - 1) + \\ &\quad + \Delta t C_1(\tau) + \dots + 1C_1(\tau + \Delta t - 1), \end{aligned} \quad (4.7)$$



4.6. ábra. Annak szemléltetése, hogy hogyan függ az időbeli korreláció,  $C(\tau)$  az ugráshoz tartozó időkülönbségtől,  $\Delta t$ -től. A korrelációt a 4.4 és a 4.5 egyenletekkel jellemzett mesterséges idősorokon mértük, a következő paraméterek mellett:  $\rho = 0.01, \tau_0 = 200, \sigma = 1000, \alpha = 0.99$ . Az ábra azt mutatja, hogy míg  $\Delta t = 1$  esetén (bal oldal) nem fedezhető fel csúcs a korrelációs függvényben, azonban megnövelve az időkülönbséget  $\Delta t = 10$  értékre (jobb ábra)  $\tau = 200$ -nál megjelenik a csúcs.

ahol  $\delta = r_{\Delta t=1}$  az egy másodperchez tartozó logaritmusos hozam érték. Megváltoztatva  $\tau$  értékét a 4.7 egyenletben azt jelenti, hogy megváltoztatjuk az egy másodperchez tartozó korrelációs függvények  $C_1$  súlyait. Mivel az eredeti adathalmaz korrelációs függvénye – (4.6) egyenlet – exponenciális lecsengésű, a maximuma,  $C_1(\tau_0)$ , fogja dominálni a (4.7) egyenletben szereplő összegzést, és mivel ennek súlya lineárisan változik  $\tau$  értékével, így  $C_{\Delta t}(\tau)$  közelítőleg lineáris lecsengést fog mutatni. (A (4.2) egyenletben szereplő normálási faktor nem változtat ezen, hiszen független  $\tau$ -tól.)

Felmerül a kérdés, hogy miként lehetséges  $\Delta t$  értékénél kisebb  $\tau$  értéket választani. A hozamok időfüggő keresztkorrelációi az  $A$  vállalat hozamának és a  $B$  vállalat  $\tau$  idővel eltolott hozamának szorzatát tartalmazzák. Mivel a hozamot a  $\Delta t$  idő-ablak definiálja, úgy tűnik, hogy  $\tau$  csak  $\Delta t$  többszöröse lehet. A probléma megoldása abban rejlik, hogy eltoljuk a  $B$  részvény hozamának kezdeti pontját  $\tau$ -val, ami a (4.3) egyenletből egyértelműen látható. Ez olyan mintha az időbeli eltolást a vállalatok árainak időfüggvényében végeznénk el, és ezáltal bármilyen, a minimális üzletkötési időnél nagyobb időtolást engedélyezünk.

Amellett érvelünk tehát, hogy az átlagolás céljából a minimális üzletkötési időnél nagyobb  $\Delta t$  időablakot kell választanunk. Ugyanakkor  $\Delta t$  értéke nem lehet túl nagy sem, mivel ekkor az átlagolás a maximum elmosódásához vezethet. Mivel az egy másodperchez tartozó korrelációs függvény szélessége legfeljebb néhány perc méretű, sokkal nagyobb időkülönbség alkalmazása azt jelentené, hogy a (4.7) egyenletben szereplő összegben többnyire csak zajból eredő tagok fognak szerepelni. Ebből az következik, hogy  $\Delta t$  optimális értéke az egy perc nagyságrendjébe esik.

### 4.2.3. Eredmények

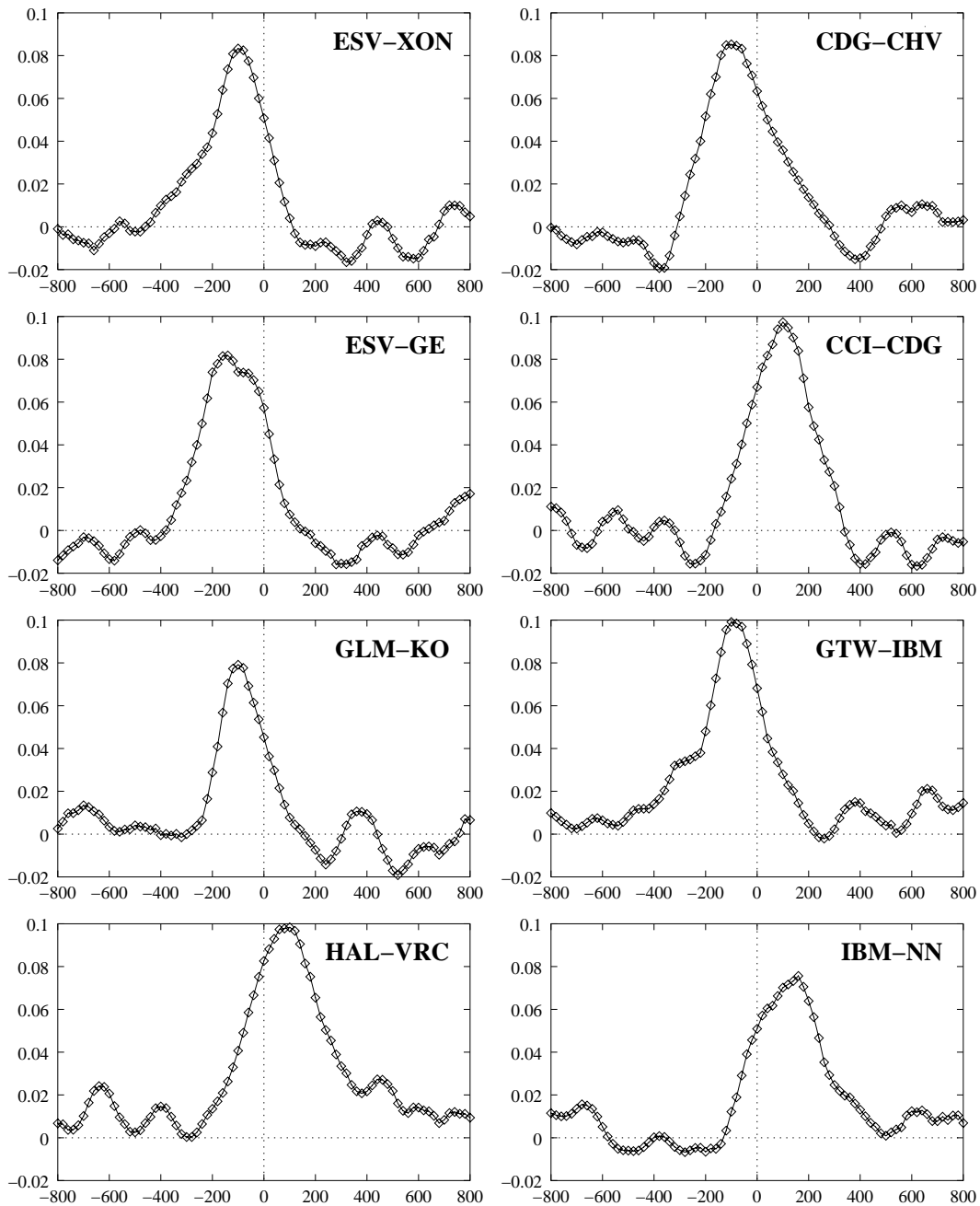
Amint azt a 4.2.1 fejezetben említettük, azon 195 vállalat közti korrelációkat vizsgáltuk, amelyeket a rendelkezésre álló 54 napban legalább 15000-szer kötöttek üzletet. Az előző fejezetben tett érveléseinkkel összhangban  $\Delta t = 100$  méretű idő-ablakot választottunk, de azt is megvizsgáltuk, hogy az eredményeink a  $50 \leq \Delta t \leq 500$  tartományban nem adnak különbséget. A maximális idő-eltolásra  $\tau = 2000$  értéket választottunk. Ez a piac hatékonysága miatt nyilvánvalóan sokkal nagyobb, mint bármilyen ésszerű korrelációs karakterisztikus idő. A fenti választásnak az volt az értelme, hogy meg tudjuk mérni a zajszintet, és össze tudjuk hasonlítani a lehetségesen fellépő effektus nagyságával.

Az adatokból nyert  $195 \cdot 194 / 2$  darab korrelációs függvény mindegyikére megmértük a maximum értékét,  $C_{max}$ , a maximum helyét,  $\tau_{max}$ , illetve a maximum és a zaj erősségének a hányadosát,  $R$ -et. A zaj erősségét a  $600 < |\tau| < 2000$  időeltolási értékek közti korrelációs értékek szórásaként definiáltuk. A vállalatok azon párjait tekintettünk, amelyekre ez a három érték meghaladott egy előírt küszöb értéket, amit  $\Delta t = 100$  esetén a következőképpen definiáltunk:  $\tau_{max} \geq 100$ ,  $C_{max} \geq 0.04$ ,  $R \geq 6.0$ .

A „mért” korrelációs függvényekre mutat példákat az a 4.7 ábra. Az első grafikon esetében például az XON (Exxon) vállalat – amely egy nagy olajipari cég – húzza az ESV-t (Enesco International), amely olajipari cégek számára végez fúrásokat. Az ábrákról az látszik, hogy minden esetben a maximális korrelációhoz tartozó eltolás a  $\tau \simeq 100$  másodperc nagyságrendjébe esik. Az effektus, noha nem túl erős ( $C_{max} \leq 0.1$ ), de a zaj szintjét jóval meghaladja, tehát mégiscsak jelentős.

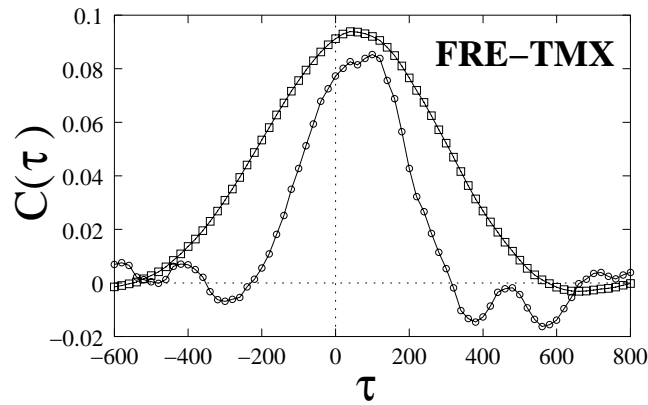
A korreláció maximális értéke minden esetben meglehetősen kicsi, átlagosan kisebb, mint 0.1, noha az egyidejű korrelációk vizsgálata ennél sokkal nagyobb értéket mutat [Mantegna, 1999, Kullmann et al., 2000a]. Ennek a jelenségnek a gyökere a hozam időkülönbségében,  $\Delta t$ -ben keresendő. Az időkülönbség növelésével növekszik az egyidejű korreláció értéke [Bonanno et al., 2001]. Ezt a jelenséget a (4.7) egyenlet is alátámasztja, hiszen a maximális korreláció súlya egyenesen arányos  $\Delta t$ -vel.

Néhány esetben azt tapasztaltuk, hogy a korreláció maximumának helye sokkal nagyobb, mint néhány perc, ami összeegyeztethetetlen a hatékony piac tulajdonsággal. Közelebbi vizsgálat azt mutatta, hogy ezekben az esetekben a korrelációs függvény maximumát az okozza, hogy mindkét vállalat idősorában szerepel egy kiemelkedően nagy hozamérték. Ezek szorzatának a korrelációhoz való hozzájárulása – az idő eltolás megfelelő értéke mellett – dominálja a korrelációs függvény maximumát. Ez azonban nem az a jelenség, amelyet mi keresünk, ezért ezeket figyelmen kívül hagytuk. Ahhoz, hogy eldöntsük, hogy a korrelációban megfigyelt csúcs egyetlen nagy hozamérték hatása-e, vagy az egyik részvénynek a másikkra tett állandó befolyásának következménye, azt is megvizsgáltuk, hogy a korrelációs függvény hogyan változik az idő függvényében, vagyis abban az esetben, ha változtatjuk a vizsgált időablakot. Tehát külön megnéztük az időfüggő korrelációkat a teljes időintervallum (54 nap) első



4.7. ábra. Néhány eredmény a vállalatok közt mért korrelációs függvényre. A vízszintes tengelyen a két idősor közti időeltolás,  $\tau$  található, míg a függőleges tengely az ehhez tartozó korreláció értékét mutatja. A maximális korrelációhoz tartozó eltolás értéke minden esetben  $\tau \simeq 100$  másodperc, és a maximális korrelációk értéke is azonos nagyságrendbe esik. Ezeknél az ábráknál  $\Delta t = 100$  időkülönbséget alkalmaztunk.

illetve a második felében is, és azt vizsgáltuk, hogy a korrelációs függvény kvalitatíve változik-e.



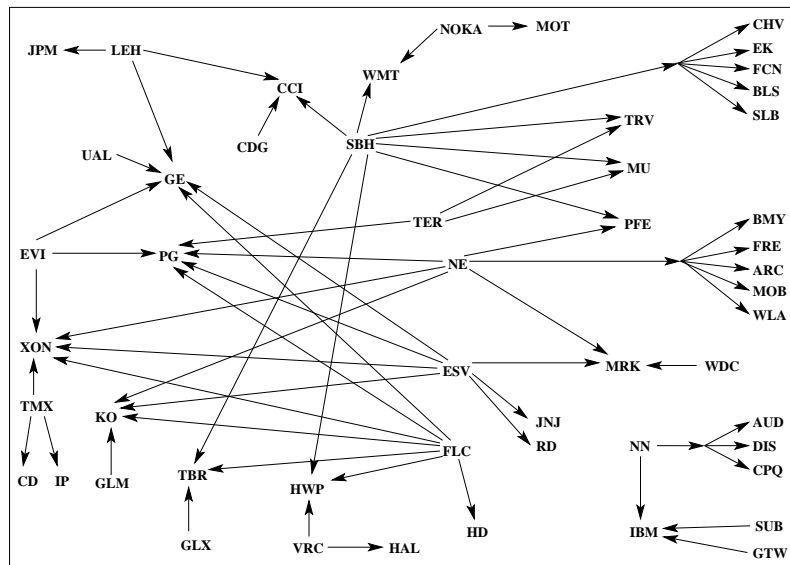
4.8. ábra. Példa egy olyan vállaltpárra, amelynél a korrelációs függvénynek aszimmetrikus csúcsa van. A körökhöz tartozó görbe  $\Delta t = 100$ -hoz tartozik, a négyzetekhez tartozó görbe pedig  $\Delta t = 500$ -hoz. A második görbe maximuma kisebb időeltolási értékhez tartozik, mivel a görbe csúcsának bal oldala magasabb értékkel rendelkezik.

Ezenkívül azt is megnéztük, hogy hogyan módosulnak az eredmények, ha a hozamhoz tartozó időkülönbséget,  $\Delta t$ -t, 100 helyett 50-nek vagy 200-nak választjuk, ugyanis előfordulhat, hogy megváltoztatva az időkülönbséget, a korreláció maximumának helye is megváltozik a (4.3) egyenletben ismerttetett átlagolás miatt. Ez akkor történhet, ha a korrelációban észlelt csúcs aszimmetrikus, amire jó példa a 4.8 ábra. Tegyük fel, hogy a függvény csúcsának bal oldala magasabb (az ehhez tartozó függvényértékek nagyobbak), mint a jobb. Ha  $\Delta t' > \Delta t$  időkülönbséget választunk, a maximum balra fog tolni, amit könnyen beláthatunk a (4.7) egyenlet segítségével. Fontos megjegyezni, hogy például az ESV-XON részvénytárhoz tartozó korrelációs függvény is aszimmetrikus, de nem a csúcsa (maximuma) közelében, ezért az időkülönbség növelésével nem fog változni maximumának helye.

Az eredmények azt mutatták (4.7 ábra), hogy a karakterisztikus idő-eltolási érték 100 másodperc körül van, ami összhangban van a hatékony piac elmélettel. Ha ugyanis az időfüggő korreláció maximális értékéhez tartozó időeltolás nagyobb lenne, mint a hozam autokorrelációjának lecsengéséhez tartozó karakterisztikus idő, lehetséges lenne az arbitrage üzletkötés.

Azt tapasztaltuk, hogy általában azok a vállalatok húzzák a másikat, amelyeknek a részvényével gyakrabban kereskednek. Ez nem meglepő, hiszen nyilvánvalóan ezek a fontosabb vállalatok, tehát sokkal valószínűbb, hogy ezek befolyásolják a kisebb vállalatokat, mint fordítva. Ugyanakkor ez csak egy általános megállapítás, hiszen létezik néhány olyan eset is, amikor az a vállalat húzza a másikat, amelyikkel ritkábban kereskednek.

A kölcsönhatások jellegéről azt a megállapítást tehetjük, hogy általában egy „kis” vállalatot sok „nagy” befolyásol, és egy „nagy” vállalat sok „kicsit” fog vonzani. A



4.9. ábra. A vállalatok közti „húzási” effektus szemléltetése. A nyilak iránya azt mutatja, hogy melyik vállalat húzza a másikat. Az ábrán azon vállalatok vannak feltüntetve, amelyek között az effektus a legjelentősebb.

kölcsönhatás ezen jellegét a 4.9 ábrán látható módon, egy irányított gráffal jellemezhetjük, ahol léteznek olyan csomópontok (a vállalatok), amelyekből sok nyíl mutat kifelé (ami azt mutatja, hogy ezt a vállalatot sok más befolyásolja), míg vannak más pontok, ahova meg sok nyíl mutat (ezek a „nagy” vállalatok, amelyek sok kisebbet vonzanak).

# Összefoglalás

Doktori munkámban a gazdaság-fizika két nagyobb témakörével foglalkoztam, a tőzsdeindex árfluktuációjának analízisével, illetve a különböző részvények közti korrelációk vizsgálatával.

1. *Karakterisztikus idők meghatározása a tőzsdeindex logaritmikus hozamának eloszlásánál*

A megfigyelések azt mutatták, hogy a logaritmikus hozamértékek eloszlása – kis hozamértékek esetén – jól jellemezhető stabil Lévy eloszlással, ezért a különböző időkülönbségekhez tartozó logaritmikus hozamértékek eloszlásai egymásba átskálázhatóak [Mantegna and Stanley, 1995]. Ugyanakkor az eloszlásban nagy hozamértékek esetén levágás figyelhető meg, emiatt a nagy időkülönbséghez tartozó logaritmikus hozamértékek eloszlása a normális eloszláshoz konvergál.

- Megmutattam, hogy a karakterisztikus idő,  $\tau_S$ , amelynél nagyobb időkülönbségekhez tartozó logaritmikus hozamértékek már nem skálázhatóak egymásba, az egy nap nagyságrendjébe esik,  $\tau_S \approx 1$  nap.

Ehhez azt vizsgáltam, hogy a különböző időkülönbségekhez tartozó hozamértékek eloszlásaira illesztett Lévy eloszlás exponense változik-e az időkülönbség függvényében, ugyanis a stabilitás feltétele, hogy az exponens állandó legyen.

- Megmutattam, hogy definiálható egy másik karakterisztikus idő,  $\tau_G$ , amely azt jellemzi, hogy mekkora az az időkülönbség, amelyhez tartozó logaritmikus hozamértékek eloszlása már Gauss-szerűnek tekinthető és hogy  $\tau_G$  az egy hónap nagyságrendjébe esik.

Az egy nap időkülönbséghez tartozó empirikus eloszlásra egy exponenciális levágású Lévy eloszlást illesztettem, meghatározva ebből az eloszlás paramétereit,  $\alpha, a_\alpha, \mu$ . Ezeket felhasználva az exponenciálisan levágott Lévy eloszlás kurtózisának analitikus alakját vizsgáltam, vagyis azt, hogy hogyan konvergál a kurtózis az időkülönbség függvényében a Gauss eloszlást jellemző  $\kappa = 3$  értékhez. Az analitikus görbe a meglehetősen szóró adatokat elfogadható módon írta le.

- Megvizsgáltam a központi határeloszlás-tétel alapján a Gauss-eloszláshoz való konvergencia feltételeit a csordamodellekben. Megállapítottam, hogy ennek kétféle eredete lehet: vagy eleve véges momentumú eloszlásra épül a modell, vagy a véges méret vezet be egy levágást. Ezek alapján rámutattam az irodalomban található ellentmondásra.

## 2. *Módosított „csordamodell”*

A Cont féle csordamodell [Cont and Bouchaud, 2000] módosított változatának segítségével vizsgáltam, hogy a szimulált árfluktuációk mennyire vannak összhangban a valós adatokon megfigyelt jelenségekkel.

A modellen két módosítást végeztem. A csoportok kialakítását a preferált növekedési modellel végeztem.

- Meghatároztam a különböző csoportok méreteloszlásának teljes időfüggő, analitikus alakját,  $\mathcal{P}(s, t)$ -t.

A modell a Barabási-féle hálózati modell leegyszerűsített változatának tekinthető és megfelel a Nobel-díjas Simon korábban csak aszimptotikusan megoldott modelljének. Az eredménynek különösen akkor van jelentősége, ha a csoportok kialakulásának, illetve pénzügyi tranzakcióknak az időskálája nem válik szét.

A mikroszkopikus modellen végzett másik módosítás, hogy az aktivitás az aktuális ár és a fundamentális ár hányadosának nemlineáris függvénye. Ennek felhasználásával a mikroszkopikus modell minőségi egyezést adott a valós árfluktuációkkal:

- Az általam bevezetett sokügynökös, egyrészvényes modellben a részvény abszolút hozamának – logaritmusos hozam abszolút értéke – korrelációja hatványfüggvény szerint cseng le. A logaritmusos hozam eloszlása szintén hatványfüggvény, és nagy hozamértékek esetén egy levágással rendelkezik, amely szintén hatványfüggvénnyel cseng le, de a stabil tartományon kívül eső exponenssel ( $\alpha = 3.36$ ).

## 3. *Különböző részvények árfluktuációi közt mért egyidejű keresztkorrelációk vizsgálata.*

A részvények csoportba rendezéséhez a szuper-paramágneses modellt [Blatt et al., 1996] használtam. A módszert úgy módosítottam, hogy ne csak ferromágneses, hanem anti-ferromágneses kölcsönhatásokat is figyelembe tudjon venni, vagyis ne csak a pozitív, hanem a negatív korrelációkat is leírja. Mind a Dow Jones mind az S&P500 tőzsdeindex részvényeire elvégeztem a csoportosítást.

- Amennyiben csak pozitív korrelációt vettem figyelembe (ferromágneses kölcsönhatás), az eredményeim megegyeztek a korábban kapott eredményekkel [Mantegna, 1999].
- Ha figyelembe vettem az antikorrrelációkat is (anti-ferromágneses kölcsönhatások), az alapállapot ( $T = 0$ ) struktúrában eltérést tapasztaltam a korábbi eredményekhez képest. Azon részvények, amelyek antikorrrelálnak szinte az összes többi részvénnyel, külön csoportot alkotnak.

#### 4. *Időfüggő kereszt-korrrelációk vizsgálata.*

A rendelkezésre álló árfolyamsorból képzett logaritmikus hozamokat stacionárius jelnek tekintve vizsgáltam az egyes részvények közötti időfüggő kereszt-korrrelációkat. A korrrelációt, mint  $\tau$  időeltolás függvényét vizsgáltam a TAQ adatbázisban elérhető összes részvénytér esetében.

- Megmutattam, hogy találhatóak olyan részvénytér-párok, amelyekre a korrrelációs függvény aszimmetrikus – a korrreláció maximuma  $\tau \neq 0$ -ban található –. Az effektus mértéke, és az időeltolás nagyságrendje (egy perc), összhangban vannak a hatékony piac feltételezéssel.
- Definiáltam és elkészítettem azt a "befolyásolási irányított hálózatot", amelynek csomópontjai a részvények. Ebben azon csomópontok vannak élekkel összekötve, amelyek esetében az időfüggő korrrelációs függvény nem szimmetrikus, és az él irányítottsága azt mutatja meg, hogy melyik részvény árfolyamváltozása követi a másikat. Megállapítottam, hogy a hálózat diszkrét részekre esik szét, soha nem fordul elő benne lánc-szerű csatolás és általában a nagy kapitalizációjú, gyakori kereskedésű részvények dominálnak. Érdekes módon kivételek is előfordulnak (pl. Nokia).

# Köszönetnyilvánítás

Ezúton szeretnék köszönetet nyilvánítani azoknak, akik hozzájárultak doktori munkám elkészítéséhez.

Mindenekelőtt Dr. Kertész Jánosnak, aki biztosította számomra a lehetőségeket szakmai fejlődésemhez, és témavezetőként irányította munkámat.

Továbbá köszönettel tartozom Kimmo Kaskinak, aki lehetőséget biztosított 2 hónapos finnországi tartózkodásomra a Helsinki University of Technology-ban, ahol együttműködésével segítette doktori munkámat. Köszönettel tartozom ezen kívül R. N. Mantegna-nak, aki kutatásaimhoz adatokat biztosított rendelkezésemre. Valamint köszönöm doktorandusz társaimnak, dr. Török Jánosnak, dr. Borda Lászlónak, Unger Tamásnak és Fazekas Sándornak, akik a nap mint nap felmerülő, elsősorban technikai jellegű problémák megoldásában segítettek.

Végül, de nem utolsó sorban köszönöm feleségemnek és édesapámnak hogy végig mellettem álltak, és kitartottak.

# A. Függelék

## Csoportméret eloszlások a mikroszkopikus modellben

Az alábbi fejezetekben néhány, a preferált csoport-kialakítási modellhez kapcsolódó mennyiség analitikus számolásával foglalkozunk, illetve megvizsgáljuk ezen mennyiségek viselkedését az aszimptotikus határesetekben. A 3.2.1 fejezetben említettük, hogy  $p = 0.5$  esetben a csoportméret eloszlás,  $\mathbf{P}(s)$  alakja megegyezik a hálózati modellre kapott eredménnyel. Hasonló egyezést tapasztalhatunk a másik két mennyiség,  $\mathcal{P}_i(s, t)$ , és  $\langle s_i \rangle(t)$  aszimptotikus viselkedésének esetében, amelyet már szintén vizsgáltak korábban [Dorogovtsev et al., 2000]

A 3.2.1 fejezetben tárgyalt, és a 3.8 ábrán szemléltetett modell leírható egy Vezér egyenlet segítségével:

$$\begin{aligned} \mathcal{P}_i(s, t) = & p \frac{(s-1)}{t-1} \mathcal{P}_i(s-1, t-1) + p \left(1 - \frac{s}{t-1}\right) \mathcal{P}_i(s, t-1) + \\ & + (1-p) \mathcal{P}_i(s, t-1) + (1-p) \Pi_{i-1}(t-1) \delta_{s,1} (1 - \delta_{i,1}), \end{aligned} \quad (\text{A.1})$$

ahol  $\mathcal{P}_i(s, t)$  annak a valószínűsége, hogy  $t$  időpontban, az  $i$ -edik csoport mérete  $s$  lesz. Az egyenlet jobb oldalán álló tagok a folyamat különböző eseteit jelenítik meg. Az első tag azt írja le, amikor az  $i$ -edik csoport mérete  $s-1$  volt, és az új elem ehhez a csoporthoz kapcsolódik. A második illetve harmadik tag annak a valószínűsége, hogy a csoport mérete  $s$  volt az előző lépésben, és vagy nem ehhez a csoporthoz kapcsolódik az új elem (második tag), vagy új csoportot alakít ki az új részecske (negyedik tag). Végül az utolsó tag annak a valószínűsége, hogy az  $i$ -edik csoport éppen az az új csoport, amit az új elem kialakít. Nyilván ez csak  $i \neq 1$  esetén igaz, mivel a modell kezdeti feltétele, hogy  $t = 1$  időpontban már létezik egy csoport egy elemmel, ezért szükséges a negyedik tagban a  $(1 - \delta_{i,1})$  kifejezés.

Az (A.1) egyenletben  $\Pi_i(t)$  annak a valószínűségét jelöli, hogy  $t$ -ben  $i$  csoport található a rendszerben:

$$\Pi_i(t) = \binom{t-1}{i-1} p^{t-1-(i-1)} (1-p)^{i-1}. \quad (\text{A.2})$$

Miután az egyes csoportok méretének eloszlását a fenti módon definiáltuk, a teljes rendszer méreteloszlása ezek átlagából számítható:

$$\mathbf{P}(s, t) = \frac{1}{t} \sum_{i=1}^t \mathcal{P}_i(s, t). \quad (\text{A.3})$$

Erről belátható, hogy aszimptotikusan,  $t \rightarrow \infty$ , egy stacionárius értékhez tart,  $\mathbf{P}(s) = \lim_{t \rightarrow \infty} \mathbf{P}(s, t)$ .

Az  $i$ -edik csoport átlagos mérete a  $t$ -edik időpillanatban szintén meghatározható  $\mathcal{P}_i(s, t)$  segítségével:

$$\langle s_i \rangle (t) = \sum_{s=1}^{t-i+1} s \mathcal{P}_i(s, t). \quad (\text{A.4})$$

Az összeg felső határa azért nem a végtelen, mivel  $\mathcal{P}_i(s, t) = 0$  ha  $s > t - i + 1$ .

## A.1. Csoportméret aszimptotikus eloszlása

Elsőként a csoportok méretének aszimptotikus eloszlását,  $\mathbf{P}(s)$ , vizsgáljuk, mivel ez az a mennyiség, amely közvetlenül kapcsolódik a gazdasági modellünkhöz. Már korábban láttuk, hogy hatványfüggvény lecsengést mutat, de most nem csak az aszimptotikus,  $s \rightarrow \infty$  esetre, hanem a teljes eloszlásra vagyunk kíváncsiak.

Ha felösszegezzük (A.1) egyenletet,  $i = 1 \dots t$ -ig, azt kapjuk, hogy:

$$t \mathbf{P}(s, t) = (t - 1 - ps) \mathbf{P}(s, t - 1) + p(s - 1) \mathbf{P}(s - 1, t - 1) + (1 - p) \delta_{s,1}, \quad (\text{A.5})$$

mivel:

$$\sum_{i=1}^t \Pi_{i-1}(t-1)(1 - \delta_{i,1}) = 1,$$

$$\sum_{i=1}^t \mathcal{P}_i(s, t - 1) = \sum_{i=1}^{t-1} \mathcal{P}_i(s, t - 1) = (t - 1) \mathbf{P}(s, t - 1).$$

Az (A.5) egyenletből látszik, hogy teljesül a stacionaritási feltétel, azaz  $\mathbf{P}(s, t \rightarrow \infty) = \mathbf{P}(s)$ . Behelyettesítve (A.5)-be a stacionárius mennyiséget, és erre rendezve az egyenletet megkapjuk  $\mathbf{P}(s)$  analitikus alakját:

$$\mathbf{P}(s) = \frac{\Gamma(s) \Gamma\left(2 + \frac{1}{p}\right)}{\Gamma\left(s + 1 + \frac{1}{p}\right)} \frac{1 - p}{1 + p} \underset{s \rightarrow \infty}{\sim} s^{-1-1/p}. \quad (\text{A.6})$$

## A.2. Az egyes csoportok méreteloszlásának analitikus megoldása

A modellben az első csoportnak kitüntetett szerepe van, mivel ez mindig tartalmaz legalább egy elemet, és emiatt ezt külön kezeljük. Az (A.1) egyenlet alakja  $i = 1$  esetén:

$$\mathcal{P}_1(s, t) = \mathcal{P}_1(s, t-1) - \frac{p}{t-1} s \mathcal{P}_1(s, t-1) + \frac{p}{t-1} (s-1) \mathcal{P}_1(s-1, t-1). \quad (\text{A.7})$$

A fenti egyenlet jobb oldalán az utolsó tag eltűnik  $s = 1$  esetén, és  $\mathcal{P}_1(1, t)$  egyszerűen kiszámítható:

$$\mathcal{P}_1(1, t) = \frac{\Gamma(t-p)}{\Gamma(t)\Gamma(1-p)}. \quad (\text{A.8})$$

Ugyanakkor  $s > 1$  esetén belátható, hogy érvényes az alábbi összefüggés:

$$\sum_{s=1}^l (-1)^{s-1} \binom{l-1}{s-1} \mathcal{P}_1(s, t) = \frac{\Gamma(t-lp)}{\Gamma(t)\Gamma(1-lp)}. \quad (\text{A.9})$$

Ha ugyanis az (A.7) egyenletet beszorozzuk a  $(-1)^{k-1} \binom{l-1}{k-1}$  taggal, és felösszegezzük  $k = 1 \dots l$ -ig, a következő kifejezést kapjuk

$$\begin{aligned} \sum_{s=1}^l (-1)^{s-1} \binom{l-1}{s-1} \mathcal{P}_1(s, t) &= \sum_{s=1}^l (-1)^{s-1} \binom{l-1}{s-1} \mathcal{P}_1(s, t-1) - \\ &- \frac{p}{t-1} \left[ \underbrace{\sum_{s=1}^l (-1)^{s-1} \binom{l-1}{s-1} s \mathcal{P}_1(s, t-1)}_x \right. \\ &\left. - \underbrace{\sum_{s=1}^l (-1)^{s-1} \binom{l-1}{s-1} (s-1) \mathcal{P}_1(s-1, t-1)}_y \right], \end{aligned} \quad (\text{A.10})$$

ahol az első tagban,  $(x)$ , leválasztjuk az összeg utolsó tagját:

$$x = \sum_{s=1}^{l-1} (-1)^{s-1} s \binom{l-1}{s-1} \mathcal{P}_1(s, t-1) + (-1)^{l-1} l \mathcal{P}_1(l, t-1). \quad (\text{A.11})$$

Figyelembe véve, hogy  $\binom{l-1}{s} = \frac{l-s}{s} \binom{l-1}{s-1}$  a második tag,  $(y)$ , az alábbi módon írható:

$$\begin{aligned} y &= \sum_{s=2}^l (-1)^{s-1} \binom{l-1}{s-1} (s-1) \mathcal{P}_1(s-1, t-1) \\ &= - \sum_{s=1}^{l-1} (-1)^{s-1} (l-s) \binom{l-1}{s-1} \mathcal{P}_1(s, t-1), \end{aligned} \quad (\text{A.12})$$

és ezáltal a két tag,  $(x)$  és  $(y)$  különbségére adódik:

$$x - y = l \sum_{s=1}^l (-1)^{s-1} \binom{l-1}{s-1} \mathcal{P}_1(s, t-1). \quad (\text{A.13})$$

Visszahelyettesítve a különbséget,  $(x - y)$ , az (A.10) egyenletbe, megkapjuk az általunk vizsgált összeg időfejlődését:

$$\sum_{s=1}^l (-1)^{s-1} \binom{l-1}{s-1} \mathcal{P}_1(s, t) = \frac{t-1-lp}{t-1} \sum_{s=1}^l (-1)^{s-1} \binom{l-1}{s-1} \mathcal{P}_1(s, t-1) \quad (\text{A.14})$$

amely megegyezik az (A.9) egyenlettel.

Az (A.9) egyenletből,  $\mathcal{P}_1(s, t)$  alakja meghatározható, amennyiben az egyenlet mindkét oldalát beszorozzuk a  $(-1)^{l-1} \binom{s-1}{l-1}$  mennyiséggel, és összegzünk  $l = 1 \dots s$ -re:

$$\mathcal{P}_1(s, t) = \sum_{l=1}^s (-1)^{l-1} \binom{s-1}{l-1} \frac{\Gamma(t-lp)}{\Gamma(t)\Gamma(1-lp)}. \quad (\text{A.15})$$

Amennyiben  $i > 1$ , az (A.1) Vezéregyenlet utolsó tagját is figyelembe kell vennünk, ami miatt az (A.9) egyenlet nem fog teljesülni. A feltételezésünk, hogy ebben az esetben a  $\mathcal{P}_i(s, t)$  valószínűség alakja az alábbi lesz:

$$\begin{aligned} \mathcal{P}_i(s, t) = & \overbrace{\sum_{l=1}^s (-1)^{l-1} \binom{s-1}{l-1} \frac{\Gamma(t-lp)}{\Gamma(t)\Gamma(1-lp)}}^{\mathcal{P}_1(s, t)} \\ & \times \left[ \sum_{b=i}^t \frac{\Gamma(b)\Gamma(1-lp)}{\Gamma(b-lp)} \binom{b-2}{i-2} p^{b-i} (1-p)^{i-1} \right]. \quad (\text{A.16}) \end{aligned}$$

Ennek érvényességét könnyen ellenőrizhetjük, ha a fenti kifejezést visszahelyettesítjük a Vezéregyenletbe. Mivel mind az (A.16) kifejezésben, mind a Vezér egyenlet jobb oldalán álló utolsó tagban szerepel az  $(1-p)^{i-1}$  kifejezés, (A Vezér egyenlet esetében ez a  $\Pi_{i-1}(t-1)$  tagban szerepel), ezzel egyszerűsíthetünk.

Az (A.1) Vezér egyenlet bal oldalán álló,  $\mathcal{P}_i(s, t)$  tagot felírhatjuk úgy, hogy az (A.16) egyenlet alapján megadott kifejezésben szereplő összegben kiválasztjuk a  $(b = t)$  tagot. Ekkor:

$$\begin{aligned} \mathcal{P}_i(s, t) = & \sum_{l=1}^s (-1)^{l-1} \binom{s-1}{l-1} \binom{t-2}{i-2} p^{t-i} \\ & + \sum_{l=1}^s (-1)^{l-1} \binom{s-1}{l-1} \frac{\Gamma(t-lp)}{\Gamma(t)\Gamma(1-lp)} \sum_{b=i}^{t-1} \frac{\Gamma(b)\Gamma(1-lp)}{\Gamma(b-lp)} \binom{b-2}{i-2} p^{b-i}, \quad (\text{A.17}) \end{aligned}$$

ahol az első tag nyilván nulla lesz, mivel  $\sum_{l=1}^{s-1} (-1)^l \binom{s-1}{l-1} = 0$ .

Ennek felhasználásával a Vezér egyenlet második és harmadik tagjának összege felírható

$$\begin{aligned} \frac{t-1-sp}{t-1} \mathcal{P}_i(s, t-1) &= (-1)^{s-1} \frac{\Gamma(t-sp)}{\Gamma(t)} \sum_{b=i}^{t-1} \frac{\Gamma(b)}{\Gamma(b-sp)} \binom{b-2}{i-2} p^{b-i} + \\ &+ \frac{t-1-sp}{t-1} \sum_{l=1}^{s-1} (-1)^{l-1} \binom{s-1}{l-1} \frac{\Gamma(t-1-lp)}{\Gamma(t-1)} \\ &\times \sum_{b=i}^{t-1} \frac{\Gamma(b)}{\Gamma(b-lp)} \binom{b-2}{i-2} p^{b-i}, \end{aligned} \quad (\text{A.18})$$

alakban, ahol az első tagnál a  $\frac{t-1-sp}{t-1}$  szorzófaktort beolvasztottuk a Gamma függvénybe.

Figyelembe véve, hogy  $(s-1) \binom{s-2}{l-1} = (s-l) \binom{s-1}{l-1}$ , a Vezér egyenlet első tagja az alábbi alakot ölti:

$$\begin{aligned} \frac{(s-1)p}{t-1} \mathcal{P}_i(s-1, t-1) &= \frac{p}{t-1} \sum_{l=1}^{s-1} (-1)^{l-1} (s-l) \binom{s-1}{l-1} \frac{\Gamma(t-1-lp)}{\Gamma(t-1)} \\ &\times \sum_{b=i}^{t-1} \frac{\Gamma(b)}{\Gamma(b-lp)} \binom{b-2}{i-2} p^{b-i}. \end{aligned} \quad (\text{A.19})$$

Összeadva (A.18) és (A.19) egyenleteket azt kapjuk, hogy:

$$\begin{aligned} &\frac{t-1-sp}{t-1} \mathcal{P}_i(s, t-1) + \frac{(s-1)p}{t-1} \mathcal{P}_i(s-1, t-1) = \\ &= (-1)^{s-1} \frac{\Gamma(t-sp)}{\Gamma(t)} \sum_{b=i}^{t-1} \frac{\Gamma(b)}{\Gamma(b-sp)} \binom{b-2}{i-2} p^{b-i} + \\ &\sum_{l=1}^{s-1} (-1)^{l-1} \binom{s-1}{l-1} \frac{\Gamma(t-lp)}{\Gamma(t)} \sum_{b=i}^{t-1} \frac{\Gamma(b)}{\Gamma(b-lp)} \frac{(b-2)}{(i-2)} p^{b-i} = \\ &= \sum_{l=1}^s (-1)^{l-1} \binom{s-1}{l-1} \frac{\Gamma(t-lp)}{\Gamma(t)} \sum_{b=i}^{t-1} \frac{\Gamma(b)}{\Gamma(b-lp)} \frac{(b-2)}{(i-2)} p^{b-i}, \end{aligned} \quad (\text{A.20})$$

amely megegyezik az (A.17) egyenlet második tagjával. Ezzel egyszerűsítve a Vezér egyenletben fennmaradó tagok az alábbi egyenlőséget adják:

$$\binom{t-2}{i-2} p^{t-i} \sum_{l=1}^s (-1)^{l-1} \binom{s-1}{l-1} = \binom{t-2}{i-2} p^{t-i} \delta_{s,1}. \quad (\text{A.21})$$

Ami könnyen belátható, hogy teljesül.

### A.3. Csoportok átlagos mérete

Helyettesítsük be az (A.16) analitikus formulát az átlagos méret definícióját megadó (A.4) egyenletbe:

$$\begin{aligned} \langle s_i \rangle (t) &= \sum_{s=1}^{t-i+1} s \sum_{l=1}^s (-1)^{l-1} \binom{s-1}{l-1} \frac{\Gamma(t-lp)}{\Gamma(t)\Gamma(1-lp)} \\ &\times \left[ \sum_{b=i}^t \frac{\Gamma(b)\Gamma(1-lp)}{\Gamma(b-lp)} \binom{b-2}{i-2} p^{b-i} (1-p)^{i-1} \right], \end{aligned} \quad (\text{A.22})$$

A két összegzés felcserélhető ( $\sum_{s=1}^{t-i+1} \sum_{l=1}^s = \sum_{l=1}^{t-i+1} \sum_{s=l}^{t-i+1}$ ), valamint teljesül az alábbi kifejezés:

$$\sum_{s=l}^{t-i+1} s \binom{s-1}{l-1} = l \binom{t-i+2}{l+1}$$

így végül a csoportok átlagos méretének időfejlődésére adódik:

$$\begin{aligned} \langle s_i \rangle (t) &= \sum_{l=1}^{t-i+1} (-1)^{l-1} l \binom{t-i+2}{l+1} \frac{\Gamma(t-lp)}{\Gamma(t)\Gamma(1-lp)} \\ &\times \sum_{b=i}^t \frac{\Gamma(b)\Gamma(1-lp)}{\Gamma(b-lp)} \binom{b-2}{i-2} p^{b-i} (1-p)^{i-1}. \end{aligned} \quad (\text{A.23})$$

### A.4. A csoportméret eloszlás, $\mathbf{P}(s, t)$ , időfüggő megoldása

Az A.1 fejezetben megmutattuk, hogy a csoportméret eloszlás aszimptotikusan egy stacionárius értékhez tart, amelyet az (A.6) egyenlet írt le. Most levezetjük a teljes, időfüggő megoldást,  $\mathbf{P}(s, t)$ .

Induljunk ki  $\mathbf{P}(s, t)$  definíciójából, az (A.3) egyenletből, és helyettesítsük be ebbe,  $\mathcal{P}_i(s, t)$  analitikus megoldását amelyet az (A.16) egyenletben határoz meg.

$$\begin{aligned} \mathbf{P}(s, t) &= \frac{1}{t} \sum_{l=1}^s (-1)^{l-1} \binom{s-1}{l-1} \frac{\Gamma(t-lp)}{\Gamma(t)\Gamma(1-lp)} \\ &\times \left[ 1 + (1-p) \sum_{i=2}^t \sum_{b=i}^t \frac{\Gamma(b)\Gamma(1-lp)}{\Gamma(b-lp)} \binom{b-2}{i-2} p^{b-i} (1-p)^{i-2} \right] \end{aligned} \quad (\text{A.24})$$

Felcserélve a két összeget,  $\sum_{i=2}^t \sum_{b=i}^t = \sum_{b=2}^t \sum_{i=2}^b$ , és figyelembe véve, hogy:

$$\sum_{b=2}^t \frac{\Gamma(b)\Gamma(1-lp)}{\Gamma(b-lp)} = \frac{\Gamma(1-lp)}{(1+lp)} \frac{\Gamma(t+1)}{\Gamma(t-lp)} - \frac{1}{1+lp},$$

végül megkapjuk a megoldást:

$$\mathbf{P}(s, t) = \underbrace{\sum_{l=1}^s (-1)^{l-1} \binom{s-1}{l-1} \left[ \frac{1-p}{1+lp} + \frac{p+lp}{1+lp} \frac{\Gamma(t-lp)}{\Gamma(t+1)\Gamma(1-lp)} \right]}_{\mathbf{P}(s, t = \infty)} \quad (\text{A.25})$$

Az egyenletben szereplő,  $\mathbf{P}(s, t = \infty)$  kifejezés az átlagos csoportméretnek az (A.6) egyenletben meghatározott stacionárius alakjával egyenlő.

## A.5. Aszimptotikus viselkedések

Eddig meghatároztuk a három legfontosabb mennyiség,  $\mathbf{P}(s, t)$ ,  $\mathcal{P}_i(s, t)$ ,  $\langle s_i \rangle(t)$  időbeli viselkedését, azonkívül egy analitikus formulát adtunk a csoportméret eloszlásra stacionárius esetben. Most azt nézzük meg, hogy hogyan viselkedik  $\mathcal{P}_i(s, t)$ ,  $\langle s_i \rangle(t)$  aszimptotikus, tehát  $t \rightarrow \infty$  esetben.

A  $\mathcal{P}_i(s, t)$  analitikus formulájában, (A.16) egyenlet, két időfüggő tag van,  $A(l, p, t)$  és  $B(i, l, p, t)$ :

$$\begin{aligned} \mathcal{P}_i(s, t) &= (1-p)^{i-1} \sum_{l=1}^s (-1)^{l-1} \binom{s-1}{l-1} \underbrace{\frac{\Gamma(t-lp)}{\Gamma(t)\Gamma(1-lp)}}_{A(l, p, t)} \\ &\times \underbrace{\left[ \sum_{b=i}^t \frac{\Gamma(b)\Gamma(1-lp)}{\Gamma(b-lp)} \binom{b-2}{i-2} p^{b-i} \right]}_{B(i, l, p, t)}. \end{aligned} \quad (\text{A.26})$$

Az első tag,  $A(l, p, t)$  aszimptotikus viselkedése egyszerűen meghatározható:

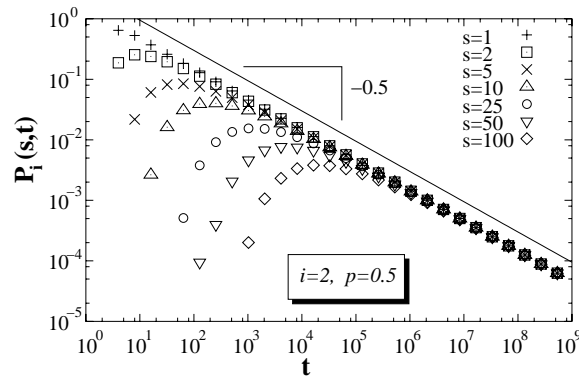
$$\lim_{t \rightarrow \infty} A(l, p, t) = \frac{1}{\Gamma(1-lp)} t^{-lp}.$$

A második tag,  $B(i, l, p, t)$  aszimptotikus esetben,  $t \gg i, l$  egy hyper-geometrikus összeghez tart [Erdélyi, 1953], és időfüggetlen lesz:

$$\lim_{t \rightarrow \infty} B(i, l, p, t) = \tilde{B}(i, l, p) = \frac{\Gamma(i)\Gamma(1-lp)}{\Gamma(i-lp)} {}_2F_1(i, i-1; i-lp; p).$$

Tehát nagy  $t$  értékekre az (A.26) egyenletben az egyetlen időfüggő tag  $t^{-lp}$  lesz, amely gyorsan lecsengő függvénye  $l$ -nek. Ezért feltehetjük, hogy  $t \gg s$  esetén a  $\mathcal{P}_i(s, t)$  mennyiségben szereplő összegben csak az első tag ( $l = 1$ ) fog nem elhanyagolható járulékot adni, tehát a különböző  $s$  értékekhez tartozó eloszlások ugyanazon  $p$  exponenssel csengenek le, amelyet igazol az A.1 ábra,

$$\lim_{t \rightarrow \infty} \mathcal{P}_i(s, t) = t^{-p} (1-p)^{i-1} \frac{\Gamma(i)}{\Gamma(i-p)} {}_2F_1(i, i-1; i-p; p) + \mathcal{O}(t^{-2p}) \quad (\text{A.27})$$

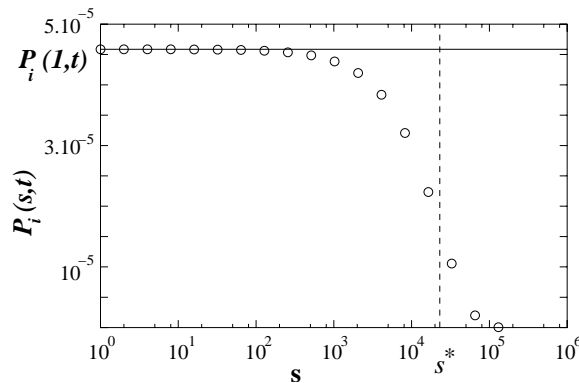


A.1. ábra.  $\mathcal{P}_i(s, t)$  aszimptotikus viselkedése. Az ábra azt szemlélteti, hogy  $t \rightarrow \infty$  határesetben a különböző  $s$  értékekhez tartozó eloszlások ugyanazzal a  $p$  exponenssel csengenek le. ( $p = 0.5$  és  $i = 2$  paramétereket választottunk.)

A fenti formula tovább egyszerűsödik, ha  $i \rightarrow \infty$ , mivel ebben az esetben teljesül, hogy  $\lim_{i \rightarrow \infty} {}_2F_1(i, i-1; i-p; p) \sim (1-p)^{1-i}$ , and  $\lim_{i \rightarrow \infty} \frac{\Gamma(i)}{\Gamma(i-p)} = i^p$ , és ezáltal az egyes csoportok méreteloszlásának aszimptotikus  $t, i \rightarrow \infty$ , viselkedésére azt kapjuk, hogy:

$$\lim_{t, i \rightarrow \infty} \mathcal{P}_i(s, t) = \left(\frac{i}{t}\right)^p \quad (\text{A.28})$$

Most térjünk rá a  $\langle s_i \rangle(t)$  mennyiség aszimptotikus vizsgálatára. Abból indulunk ki, hogy  $s \ll t$  esetén, a csoportméret eloszlás,  $\mathcal{P}_i(s, t)$ , leírható az összeg első tagjával, mint azt az (A.27) egyenletben demonstráltuk, ugyanakkor nagyobb  $s$  értékekre a mennyiség gyors lecsengést mutat, ahogy az A.2 ábra is mutatja. Definiálhatunk tehát



A.2. ábra. Az egyes csoportok méreteloszlása, mint a csoportméret függvénye,  $t \rightarrow \infty$  esetén. ( $t = 10^9$  értéket választottunk.) Kis csoportméret értékeknél –  $s$  kisebb, mint egy bizonyos küszöbérték – a valószínűség jól közelíthető a  $\mathcal{P}_i(s = 1, t)$  mennyiséggel, de ha  $s$  meghaladja a küszöb értékét,  $\mathcal{P}_i(s, t)$  gyors lecsengést mutat.

egy levágási értéket,  $s^*$ , amelynek segítségével feltehetjük, hogy az átlagos csoport

méret (A.4)-ben meghatározott alakja az alábbi alakot ölti:

$$\langle s_i \rangle(t) \approx \sum_{s=1}^{s^*} s \mathcal{P}_i(s, t) = \mathcal{P}_i(1, t) \frac{s^*(s^* + 1)}{2}. \quad (\text{A.29})$$

A levágási küszöbérték sokféle módon választható, mi a  $\mathcal{P}_i(s, t)$  mennyiség inflexiós pontjaként definiáltuk, tehát:

$$\begin{aligned} s^* &= t^p \frac{\tilde{B}(i, 3, p)}{\tilde{B}(i, 4, p)} \frac{\Gamma(1 - 4p)}{\Gamma(1 - 3p)} + 2 + \mathcal{O}(t^{-p}) \\ &= t^p \frac{\Gamma(i - 4p)}{\Gamma(i - 3p)} \frac{{}_2F_1(i, i - 1; i - 3p; p)}{{}_2F_1(i, i - 1; i - 4p; p)} + 2 + \mathcal{O}(t^{-p}). \end{aligned} \quad (\text{A.30})$$

Behelyettesítve (A.29)-be adódik:

$$\begin{aligned} \langle s_i \rangle(t) &\simeq t^p (1 - p)^{i-1} \frac{\Gamma(i)}{\Gamma(i - p)} {}_2F_1(i, i - 1; i - p; p) \\ &\quad \times \left[ \frac{\Gamma(i - 4p)}{\Gamma(i - 3p)} \frac{{}_2F_1(i, i - 1; i - 3p; p)}{{}_2F_1(i, i - 1; i - 4p; p)} \right]^2. \end{aligned} \quad (\text{A.31})$$

Hasonlóan  $\mathcal{P}_i(s, t)$  aszimptotikus vizsgálatához, ez a mennyiség is jelentősen egyszerűsödik  $i \rightarrow \infty$  esetben, mivel ekkor az alábbi határértékeket vehetjük figyelembe:  $\lim_{i \rightarrow \infty} \frac{\Gamma(i)}{\Gamma(i - p)} = i^p$ ,  $\lim_{i \rightarrow \infty} \frac{\Gamma(i - 4p)}{\Gamma(i - 3p)} = i^{-p}$ ,  $\lim_{i \rightarrow \infty} {}_2F_1(i, i - 1; i - 3p; p) = \lim_{i \rightarrow \infty} {}_2F_1(i, i - 1; i - 4p; p)$ ,  $\lim_{i \rightarrow \infty} {}_2F_1(i, i - 1; i - p; p) \sim (1 - p)^{1-i}$ , amelyek felhasználásával végül eljutunk az átlagos csoportméret végső, aszimptotikus alakjához:

$$\lim_{t, i \rightarrow \infty} \langle s_i \rangle(t) = \left( \frac{t}{i} \right)^p \quad (\text{A.32})$$

## B. Függelék

# Valószínűségi változók összegének szórásnégyzete és a változók autokorrelációjának exponense közti összefüggés levezetése

Vizsgáljuk meg, hogy  $n$  valószínűségi változó összegének szórása és a változó autokorrelációja milyen viszonyban állnak egymással. Definíció alapján az összeg szórásnégyzete az alábbi alakba írható.

$$\begin{aligned}
 D^2\left(\sum_{i=1}^n x_i\right) &\equiv \overline{\left[(x_1 + \dots + x_n) - \overline{(x_1 + \dots + x_n)}\right]^2} = \overline{(x_1 - \bar{x}_1)^2} + \dots + \overline{(x_n - \bar{x}_n)^2} + \\
 &\quad + 2\overline{(x_1 - \bar{x}_1)(x_2 - \bar{x}_2)} + 2\overline{(x_1 - \bar{x}_1)(x_3 - \bar{x}_3)} + \dots + 2\overline{(x_1 - \bar{x}_1)(x_n - \bar{x}_n)} + \\
 &\quad + 2\overline{(x_2 - \bar{x}_2)(x_3 - \bar{x}_3)} + \dots + 2\overline{(x_2 - \bar{x}_2)(x_n - \bar{x}_n)} + \\
 &\quad + \dots + \\
 &\quad + 2\overline{(x_{n-1} - \bar{x}_{n-1})(x_n - \bar{x}_n)}.
 \end{aligned}$$

A jobb oldalon szereplő első  $n$  tag a valószínűségi változó szórásnégyzete,  $D^2(1)$ , míg a többi tag a különböző időkülönbségekhez tartozó autokovarianciákat jelöli. Ennek figyelembevételével az összeg szórása a következőképpen írható fel:

$$\begin{aligned}
 D^2\left(\sum_{i=1}^n x_i\right) &= nD^2(1) + 2D^2(1)[(n-1)C(1) + (n-2)C(2) + \dots + C(n-1)] = \\
 &\quad D^2(1) \left[ n + 2 \sum_{m=1}^{n-1} (n-m)C(m) \right]. \tag{B.1}
 \end{aligned}$$

Tegyük fel, hogy  $D^2(1) = 1$ , valamint, hogy az autokorreláció hatványfüggvény lecsengésű, azaz  $C(m) \sim m^{-\kappa}$ .

$$D^2\left(\sum_{i=1}^n x_i\right) = n + 2 \sum_{m=1}^{n-1} (n-m) \frac{1}{m^\kappa} \quad (\text{B.2})$$

Azt is feltehetjük, hogy az összeadott változók száma nagy, így  $n-1 \simeq n$  közelítéssel élhetünk. Ekkor a (B.2) egyenlet második tagja a következőképpen írható:

$$\sum_{m=1}^n (n-m) \frac{1}{m^\kappa} \simeq \left\{ \tilde{m} = \frac{m}{n} \right\} = n^{1-\kappa} \sum_{\tilde{m}=1/n}^1 \frac{1-\tilde{m}}{\tilde{m}^\kappa}. \quad (\text{B.3})$$

Itt egy változócsere-t hajtottunk végre, bevezetve az  $\tilde{m} = m/n$  mennyiséget. Amennyiben  $n$  értéke nagy, az összegzés lépéseleme,  $\Delta\tilde{m} = 1/n$  kicsi, így az összegzés helyettesíthető integrálással. Tehát az összeg szórásnégyzete a következő alakot ölti:

$$D^2\left(\sum_{i=1}^n x_i\right) = n + 2n^{1-\kappa} \int_{1/n}^1 \frac{1-\tilde{m}}{m^\kappa} n d\tilde{m} = n + 2n^{2-\kappa} \left[ \left( \frac{1}{1-\kappa} - \frac{n^{\kappa-1}}{1-\kappa} \right) - \left( \frac{1}{2-\kappa} - \frac{n^{\kappa-2}}{2-\kappa} \right) \right]. \quad (\text{B.4})$$

Az egyenlet átrendezésével a valószínűségi változók összegének szórásnégyzete – hatványfüggvény lecsengésű autokorreláció feltételezése mellett – az alábbi alakot ölti:

$$D^2\left(\sum_{i=1}^n x_i\right) = \frac{2}{(1-\kappa)(2-\kappa)} n^{2-\kappa} + \frac{\kappa+1}{\kappa-1} n + \frac{2}{2-\kappa}. \quad (\text{B.5})$$

A fenti egyenletben az utolsó tag, mivel  $n$ -ben nulladrendű, elhanyagolható a másik két tag mellett. Amennyiben a korreláció gyors lecsengésű, azaz  $2 \geq \kappa \geq 1$ , a második tag fog dominálni, így a szórásnégyzetre fennáll, hogy  $D^2(n) \sim n$ . Abban az esetben viszont, ha a korreláció lassú lecsengésű,  $\kappa < 1$  a fenti kifejezés első tagja dominálja az összeget, és így a valószínűségi változók összegének szórásnégyzete  $\kappa$  függvénye lesz, azaz  $D^2(n) \sim n^{2-\kappa}$ . Ezzel beláttuk a (3.15) egyenletben szereplő kifejezést.

## C. Függelék

### Vállalatok neve, szimbóluma, és rövid leírása

<b>Szimbólum</b>	<b>Név</b>	<b>Leírás</b>
AA	Alcoa Inc	Aluminium-ipar
ALD	AlliedSignal Inc.	Nyersanyag gyártás
ARC	Atlantic Richfield Co.	Petróleum finomítás
AUD	Automatic Data Processing	Információ szolgáltatás
AXP	American Express	Utazási szolgáltatások
BA	Boeing Co.	Repülőgépgyártás
BLS	Bellsouth Corp.	Telefon
BMJ	Bristol-Myers Squibb Co.	Gyógyszerkészítés
BS	Bethlehem Steel	Acélipar
CAT	Caterpillar Inc.	Gépgyártás
CCI	Citicorp	Bank
CD	Cendant Corp.	Utazási szolgáltatások
CDG	Cliffs Drilling Co.	Nemzetközi olajfúró vállalat
CHV	Chevron Corp.	Vegyipar, petróleum finomítás
CPQ	Compaq Computer Corp.	Elektronika, számítástechnika
DD	duPont(EI)deNemours	Olajipar
DIS	Walt Disney Co.	Szórakoztatóipar
EK	Eastman Kodak Co.	Fényképezés
ESV	Ensco International Inc.	Olajfúrás, olajkút
EVI	Energy Ventures Inc.	Olaj- és gázipar
FCN	First Chicago NBD Corp.	Bank
FLC	Falcon Drilling Co. Inc.	Tengeri olajfúrás
GE	General Electric Co.	Elektronika, gépipar
GLM	Global Marine Inc.	Olajfúrás
GM	General Motors	Autogyártás
GLX	Glaxo Wellcome Plc.	Gyógyszeripar
GTW	Gateway 2000 Inc.	Elektronika, számítástechnika
HAL	Halliburton Co.	Olajipar
HD	Home Depot Inc.	Háztartási cikkek

Szimbólum	Név	Leírás
HWP	Hewlett-Packard Co.	Számítástechnika
IBM	International Business Machines Corp.	Számítástechnika
IP	International Paper Co.	Papíripar
JNJ	Johnson & Johnson	Egészségügy
JPM	Morgan J.P. Co. Inc.	Bank
KO	Coca-Cola Co.	Üdítőital gyártás
LEH	Lehman Brothers Holdings	Pénzügyi szolgáltatások
MCD	McDonald's Corp.	Gyorsétterem
MMM	Minnesota Mining/Mfg	Vegyipar
MO	Philip Morris Cos.	Cigaretta
MOB	Mobil Corp.	Petróleum finomítás
MOT	Motorola Inc.	Félvezető technológia
MRK	Merck & Co Inc.	Gyógyszeripar
MU	Micron Technology Inc.	Félvezető technológia
NE	Noble Drilling Corp.	Olajfúrás
NN	Newbridge Networks Corp.	Telefon
NOKA	Nokia Corp.	Mobiltelefon
PFE	Pfizer Inc.	Gyógyszeripar
PG	Procter & Gamble Co.	Háztartási vegyiáru
RD	Royal Dutch Petroleum Comp.	Petróleum finomítás
SBH	Smithkline Plc	Gyógyszeripar
SLB	Schlumberger Limited LTD	Olaj-, gázipar
SUB	Summit Bank Corp.	Bank
T	AT&T Corp.	Telekommunikáció
TBR	Telecomunicacoes Brasileiras S.A.	Telekommunikáció
TER	Teradyne Inc.	Elektronikai berendezések
TMX	Telefonos de Mexico	Telefon
TRV	Travelers Group Inc.	Biztosítás
TX	Texaco Inc.	Olajipar
UAL	UAL Corp.	Légiközlekedés
UK	Union Carbide	Vegyipar
UTX	United Technologies	Gépgyártás
VRC	Varco International Inc.	Olaj-, gázipar
WDC	Western Digital Corp.	Számítástechnika
WLA	Warner Lambert Co.	Gyógyszeripar
WMT	Wal-Mart Stores Inc.	Vegyeskereskedés
XON	Exxon Corp.	Petróleum finomítás

# Irodalomjegyzék

- [Akgiray, 1989] Akgiray, V. *Conditional Heteroscedasticity in Time Series of Stock Returns: Evidence and Forecasts*. Journal of Business **62**, pp. 55–80 (1989).
- [Albert et al., 1999] Albert, R., Jeong, H., and Barabási, A. L. *Diameter of The World-Wide Web*. Nature **401**, pp. 130 (1999).
- [Arnéodo et al., 1998] Arnéodo, A., Muzy, J.-F., and Sornette, D. *"Direct" causal cascade in the stock market*. Eur. Phys. J. B **2**, pp. 277–282 (1998).
- [Bachelier, 1900] Bachelier, L. *Théorie de la spéculation*. Annales Scientifiques de l'Ecole Normale Supérieure **III:17**, pp. 21–86 (1900).
- [Barabási and Albert, 1999] Barabási, A. L. and Albert, R. *Emergence of Scaling in Random Networks*. Science **286**, pp. 509 (1999).
- [Barabási et al., 1999] Barabási, A. L., Albert, R., and Jeong, H. *Mean-field theory for scale-free random networks*. Physica A **272**, pp. 173–187 (1999).
- [Berg and Neuhaus, 1992] Berg, B. A. and Neuhaus, T. *Multicanonical Ensemble: A New Approach to Simulate First-Order Phase Transitions*. Phys. Rev. Lett. **68**, pp. 9–12 (1992).
- [Black and Scholes, 1973] Black, F. and Scholes, M. *The pricing of options and corporate liabilities*. Journal of Political Economy **81**, pp. 637–654 (1973).
- [Blatt et al., 1996] Blatt, M., Wiseman, S., and Domany, E. *Superparamagnetic Clustering of Data*. Phys. Rev. Lett. **76**, No. 18, pp. 3251 (1996).
- [Bollerslev, 1986] Bollerslev, T. *Generalized Autoregressive Conditional Heteroscedasticity*. Journal of Econometrics **31**, No. 1, pp. 307–327 (1986).
- [Bollerslev and Chou, 1992] Bollerslev, T. and Chou, R. C. *ARCH modeling in finance*. Journal of Econometrics **52**, pp. 5 (1992).
- [Bonanno et al., 2001] Bonanno, G., Lillo, F., and Mantegna, R. N. *High-frequency Cross-correlation in a Set of Stocks*. Quantitative Finance **1**, pp. 96–104 (2001).

- [Bonanno et al., 2000] Bonanno, G., Vandewalle, N., and Mantegna, R. N. *Taxonomy of Stock Market Indices*. Phys. Rev. E **62**, pp. R7615–R7618 (2000).
- [Bouchaud and Potters, 1997] Bouchaud, J.-P. and Potters, M. *Théories des Risques Financiers*. Aléa-Saclay, Paris (1997).
- [Chang and Stauffer, 1999] Chang, I. and Stauffer, D. *Fundamental judgement in Cond-Bouchaud Herding model of market fluctuations*. Physica A **264**, pp. 294 (1999).
- [Chowdhury and Stauffer, 1999] Chowdhury, D. and Stauffer, S. *A generalized spin model of financial markets*. Eur. Phys. J. B **8**, pp. 477 (1999).
- [Cizeau et al., 1997] Cizeau, P., Liu, Y., Meyer, M., Peng, C.-K., and Stanley, H. E. *Volatility distribution in the S&P500 Stock Index*. Physica A **245**, pp. 441 (1997).
- [Cont, 1997] Cont, R. (1997). *Scaling properties of intraday price changes: a study of S&P futures*. cond-mat/9705075.
- [Cont and Bouchaud, 2000] Cont, R. and Bouchaud, J.-P. *Herd behavior and aggregate fluctuations in financial markets*. Journal of Macroeconomic Dynamics **4**, No. 2, pp. 170 (2000).
- [Cont et al., 1997] Cont, R., Potters, M., and Bouchaud, J.-P. *Scaling in stock market data: stable laws and beyond*. CNRS Workshop on Scale Invariance (1997).
- [Dorogovtsev et al., 2000] Dorogovtsev, S. N., Mendes, J. F. F., and Shamukin, A. N. *Structure of growing networks with preferential linking*. Phys. Rev. Lett. **85**, pp. 4633–4636 (2000).
- [Elton and Gruber, 1995] Elton, E. J. and Gruber, M. J. *Modern Portfolio Theory and Investment Analysis*. J. Wiley and Sons, New York (1995).
- [Engle, 1982] Engle, R. F. *Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation*. Econometrica **50**, No. 4, pp. 987–1007 (1982).
- [Erdélyi, 1953] Erdélyi, A. *Higher Transcendental Functions*. McGraw-Hill Book Co., INC. (1953).
- [Erdős and Rényi, 1960] Erdős, P. and Rényi, A. *On the evolution of random graphs*. Publications of the Mathematical Institute of the Hungarian Academy of Sciences **5**, pp. 17–61 (1960).
- [Faloutsos et al., 1999] Faloutsos, M., Faloutsos, P., and Faloutsos, C. *On Power-Law Relationship of the Internet Topology*. Comput. Comm. Rev. **29**, pp. 251 (1999).

- [Fama, 1970] Fama, E. F. *Efficient Capital Markets: A Review of Theory and Empirical Work*. J. Finance **25**, pp. 383–417 (1970).
- [Feller, 1968] Feller, W. *An Introduction to Probability Theory and Its Applications 2*. Wiley (1968).
- [Furth, 1917] Furth, R. Annalen der Physik **53**, pp. 177 (1917).
- [Giada and Marsili, 2001] Giada, L. and Marsili, M. *Data clustering and noise undressing of correlation matrices*. Phys. Rev. E. **63**, pp. 061101 (2001).
- [Gopikrishnan et al., 1998] Gopikrishnan, P., Meyer, M., Amaral, L. A. N., and Stanley, H. E. *Inverse cubic law for the distribution of stock price variations*. Eur. Phys. J. B **3**, pp. 319 (1998).
- [Gopikrishnan et al., 1999] Gopikrishnan, P., Plerou, V., and Stanley, H. E. *Scaling of the distribution of fluctuations of financial market indices*. Phys. Rev. E **60**, pp. 5305 (1999).
- [Hill, 1975] Hill, B. *A Simple General Approach to Inference About the Tail of a Distribution*. Annals of Statistics **3**, pp. 1163–1174 (1975).
- [Hsieh, 1991] Hsieh, D. A. *Chaos and non-linear dynamics: application to financial markets*. Journal of Finance **46**, pp. 1839 (1991).
- [Kertész and Kondor, 1996] Kertész, J. and Kondor, I., editors. *Advances in Computer Simulation*. Springer, Berlin (1996).
- [Kullmann and Kertész, 2001a] Kullmann, L. and Kertész, J. *Crossover to Gaussian behavior in herding market models*. International Journal of Modern Physics C **12**, pp. 1 (2001).
- [Kullmann and Kertész, 2001b] Kullmann, L. and Kertész, J. *Preferential growth: Exact solution of the time-dependent distributions*. Phys. Rev. E **63**, pp. 051112 (2001).
- [Kullmann and Kertész, 2001c] Kullmann, L. and Kertész, J. *Preferential growth: solution and application to modeling stock market*. Physica A **299**, pp. 121 (2001).
- [Kullmann et al., 2002] Kullmann, L., Kertész, J., and Kaski, K. *Time dependent cross correlations between different stock returns: A directed network of influence*. Phys. Rev. E **66**, pp. 026125 (2002).
- [Kullmann et al., 2000a] Kullmann, L., Kertész, J., and Mantegna, R. N. *Identification of clusters of companies in stock indices via Potts super-paramagnetic transitions*. Physica A **287**, pp. 412–419 (2000).

- [Kullmann et al., 1999] Kullmann, L., Töyli, J., Kertész, J., Kanto, A., and Kaski, K. *Characteristic times in stock market indices*. Physica A **269**, pp. 98–110 (1999).
- [Kullmann et al., 2000b] Kullmann, L., Töyli, J., Kertész, J., Kanto, A., and Kaski, K. *Break-down of scaling and convergence to Gaussian distribution in stock market data*. IJTAF **3**, No. 3, pp. 371–373 (2000).
- [Laloux et al., 1998] Laloux, L., Bouchaud, P. C. J.-P., and Potters, M. *Noise Dressing of Financial Correlation Matrices*. Phys. Rev. Lett. **83**, pp. 1467–1470 (1998).
- [Lillo and Mantegna, 2000a] Lillo, F. and Mantegna, R. N. *Symmetry alteration of ensemble return distribution in crash and rally days of financial markets*. Eur. Phys. J. B **15**, pp. 603–606 (2000).
- [Lillo and Mantegna, 2000b] Lillo, F. and Mantegna, R. N. *Variety and volatility in financial markets*. Phys. Rev. E **62**, pp. 6126–6134 (2000).
- [Liu et al., 1999] Liu, Y., Gopikrishnan, P., Cizeau, P., Meyer, M., Peng, C., and Stanley, H. E. *Statistical properties of the volatility of price fluctuations*. Phys. Rev. E **60**, pp. 1390 (1999).
- [Lux and Marchesi, 1999] Lux, T. and Marchesi, M. *Scaling and criticality in stochastic multi-agent model of a financial market*. Nature **397**, pp. 498–500 (1999).
- [Mandelbrot, 1963] Mandelbrot, B. *The variation of certain speculative prices*. Journal of Business **26**, pp. 394–419 (1963).
- [Mantegna and Stanley, 2000] Mantegna, R. and Stanley, H. *Econophysics*. Cambridge UP (2000).
- [Mantegna, 1999] Mantegna, R. N. *Hierarchical structure in financial markets*. Eur. Phys. J. B. **11**, pp. 193 (1999).
- [Mantegna and Stanley, 1994] Mantegna, R. N. and Stanley, H. E. *Stochastic Properties with Ultraslow Convergence to a Gaussian: The Truncated Levy Flight*. Phys. Rev. Lett. **73**, No. 22, pp. 2946–2949 (1994).
- [Mantegna and Stanley, 1995] Mantegna, R. N. and Stanley, H. E. *Scaling behaviour in the dynamics of an economic index*. Nature **376** (1995).
- [Mantegna and Stanley, 1998] Mantegna, R. N. and Stanley, H. E. *Modeling of financial data: Comparison of the truncated Lévy flight and the ARCH(1) and GARCH(1,1) processes*. Physica A **254**, pp. 77–84 (1998).
- [Maslov, 2000] Maslov, S. *Simple model of a limit order-driven market*. Physica A **278**, pp. 571–578 (2000).

- [Mehta, 1995] Mehta, M. *Random Matrices*. Academic Press, New York (1995).
- [Pasquini and Serva, 2000] Pasquini, M. and Serva, M. *Clustering of volatility as a multiscale phenomenon*. Eur. Phys. J. B **16**, pp. 195–201 (2000).
- [Peng et al., 1994] Peng, C.-K., Buldyrev, S. V., Havlin, S., Simons, M., and Goldberger, H. E. S. A. L. *Mosaic organization of DNA nucleotides*. Phys. Rev. E **49**, No. 2, pp. 1685–1689 (1994).
- [Plerou et al., 2000] Plerou, V., Gopikrishnan, P., Amaral, L. A., Gabaix, X., and Stanley, H. E. *Economic fluctuation and anomalous diffusion*. Phys. Rev. E **62**, pp. R3023 (2000).
- [Plerou et al., 1999] Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N., and Stanley, H. E. *Universal and Nonuniversal Properties of Cross Correlations in Financial Time Series*. Phys. Rev. Lett. **83**, No. 7, pp. 1471 (1999).
- [Potters et al., 1998] Potters, M., Cont, R., and Bouchaud, J.-P. *Financial markets as adaptive system*. Europhys. Lett. **41**, No. 3, pp. 239–244 (1998).
- [Rammal et al., 1986] Rammal, R., Toulouse, G., and Virasoro, M. A. *Ultrametricity for physicists*. Rev. Mod. Phys. **58**, No. 3, pp. 765–788 (1986).
- [Redner, 1998] Redner, S. *How Popular is Your Paper? An Empirical Study of the Citation Distribution*. Eur. Phys. J. B **4**, pp. 131 (1998).
- [Sengupta and Mitra, 1997] Sengupta, A. M. and Mitra, P. P. *Distribution of Singular Values for Some Random Matrices*. Phys. Rev. E. Rapid Communication, pp. 3389–3395 (1997).
- [Simon, 1955] Simon, H. A. *On a class of skew distribution functions*. Biometrika **42**, pp. 425 (1955).
- [Späth, 1978] Späth, H. *Spline-Algorithmen zur Konstruktion glatter Kurven und Flächen*. Oldenbourg, München (1978).
- [Stauffer and Aharony, 1994] Stauffer, D. and Aharony, A. *Introduction to Percolation Theory, 2nd edition*. Taylor and Francis (1994).
- [Stauffer and Penna, 1998] Stauffer, D. and Penna, T. J. P. *Crossover in the Cont-Bouchaud percolation model for market fluctuations*. Physica A **256**, pp. 284–290 (1998).
- [Stauffer and Sornette, 1999] Stauffer, D. and Sornette, D. *Self-organized percolation model for stock market fluctuations*. Physica A **271**, pp. 496 (1999).

- [Swendsen and Wang, 1987] Swendsen, R. H. and Wang, J.-S. *Nonuniversal critical dynamics in Monte Carlo simulations*. Phys. Rev. Lett. **58**, pp. 86–88 (1987).
- [Weiss, 1994] Weiss, G. H. *Aspects and Applications of the Random Walk*. North-Holland, Amsterdam (1994).
- [Weron, 2001] Weron, R. *Levy stable distribution revisited: tail index  $> 2$  does not exclude the Levy stable regime*. International Journal of Modern Physics C **12**, No. 2, pp. 209–223 (2001).
- [Zhang, 1999] Zhang, Y.-C. *Towards a theory of marginally efficient markets*. Physica A **269**, pp. 30 (1999).
- [Zumbach, 1998] Zumbach, G. O. *Considering Time as the random variable: the first hitting time*. Neutral Network World **8**, pp. 243–253 (1998).

# Publikációs lista

1. L. Kullmann, J. Töyli, J. Kertész, A. Kanto and K. Kaski, *Characteristic times in stock market indices*, Physica A **269**, pp. 98-110, 1999
2. L. Kullmann, J. Kertész, J. Töyli, K. Kaski, A. Kanto, *Break-down of scaling and convergence to gaussian distribution in stock market data*, IJTAF **3**, No. 3 pp. 371-373, 2000
3. L. Kullmann, J. Kertész and R.N. Mantegna, *Identification of clusters of companies in stock indices via Potts super-paramagnetic transitions*, Physica A **287** pp. 412-419, 2000
4. L. Kullmann, J. Kertész, *Preferential growth: Exact solution of the time-dependent distributions*, Phys. Rev. E **63**, pp. 051112, 2001
5. L. Kullmann, J. Kertész, *Preferential growth: Solution and application to modeling stock market*, Physica A **299** pp. 201-126, 2001
6. L. Kullmann, J. Kertész, *Crossover to Gaussian behavior in herding market models*, Int. J. Mod. Phys. C **12**, No. 8, 2001
7. L. Kullmann, J. Kertész, K. Kaski, *Time dependent cross correlations between different stock returns: A directed network of influence*, Phys. Rev. E **66** pp. 026125, 2002
8. J. Kertész, L. Kullmann, A.G. Zawadowski, R. Karádi and K. Kaski, *Correlations and response: absence of detailed balance on the stock market* Physica A, (nyomdában)