

# DOKTORI (PhD) ÉRTEKEZÉS

## Kockázati mérések és mértékek instabilitása

Szerző: Nagy Gábor



Témavezető: Dr. Gáll József Mihály

Debreceni Egyetem  
Közgazdaságtudományi Doktori Iskola  
Versenyképesség, Globalizáció, Regionalitás program  
Debrecen, 2015.05.04.

**A doktori értekezés betétlapja**

**Kockázati mérések és mértékek instabilitása**

Értekezés a doktori (PhD) fokozat megszerzése érdekében  
a közgazdaságtudomány tudományágban

Írta: Nagy Gábor okleveles közgazdász

Készült a Debreceni Egyetem Közgazdaságtudomány Doktori Iskolája  
(Versenyképesség, globalizáció, regionalitás programja) keretében

Témavezető: Dr. Gáll József Mihály

A doktori szigorlati bizottság:

elnök: Dr. ....  
tagok: Dr. ....  
Dr. ....

A doktori szigorlat időpontja: 2015... ..

Az értekezés bírálói:

Dr. ....  
Dr. ....  
Dr. ....

A bírálóbizottság:

elnök: Dr. ....  
tagok: Dr. ....  
Dr. ....  
Dr. ....  
Dr. ....

Az értekezés védésének időpontja: 2015... ..

Nyilatkozat

Alulírott Nagy Gábor nyilatkozom, hogy

- a) értekezésemet korábban más intézményben nem nyújtottam be és azt nem utasították el;
- b) nem állok doktori fokozat visszavonására irányuló eljárás alatt, illetve 5 éven belül nem vontak vissza tőlem korábban odaítélt doktori fokozatot;
- c) a disszertáció önálló munkám, az irodalmi hivatkozások egyértelműek és teljesek.

Kelt: Debrecen, 2015. ....

.....  
Nagy Gábor

## Tartalomjegyzék

Tartalomjegyzék.....	3
1. Köszönetnyilvánítás .....	5
2. Bevezetés.....	6
3. Működési kockázatok tőkeszámításának lehetőségei .....	15
4. Basel II keretrendszer - AMA .....	17
4.1. Belső adatok .....	17
4.1.1. Dátum választás.....	18
4.1.2. Időtáv választás .....	18
4.1.3. Modellezett mennyiség megválasztása .....	19
4.1.4. Csoportos veszteségek.....	19
4.1.5. Adatgyűjtési küszöb .....	19
4.2. Külső adatok.....	19
4.3. Stilizált tények.....	21
4.4. Kockázati filozófia .....	22
5. A tőkekövetelmény-számítás filozófiája.....	23
5.1. A tőkekövetelmény LDA esetén .....	23
5.2. Tőkekövetelmény számítás egyéb módszerei .....	30
5.2.1. Közelítéssel módszer.....	30
5.2.2. Single Loss Approximation.....	32
6. Az operációs kockázat becslése, és kapcsolódó problémák.....	33
6.1. A mintáról .....	35
6.1.1. Az optimális mintaelemszámról.....	36
6.1.2. Single Loss Approximation és mintanagyság .....	39
6.2. A gyakoriság eloszlásról .....	41
6.2.1. Tipikus gyakoriság eloszlások .....	41
6.2.2. Gyakoriság eloszlás választás .....	42
6.2.3. Single Loss Approximation és gyakoriság eloszlás választás.....	47
6.3. A veszteségeloszlás .....	47
6.3.1. Gyakran alkalmazott veszteségeloszlások és illesztésük .....	47
6.3.2. Single Loss Approximation és az egyedi veszteségeloszlás .....	50
6.4. Adatgyűjtési küszöb .....	54
6.4.1. Adatgyűjtési küszöb hatása a tőkére .....	54
6.4.2. Adatgyűjtési küszöb és az egyedi veszteségeloszlás paramétereinek Maximum Likelihood becslése .....	56
6.5. Azonos mintaelemek .....	60
7. Tőkekövetelmény-számítás gyakorlati megvalósítása .....	60
7.1. Az összetett eloszlás (LDA) és a tőkekövetelmény .....	60
7.1.1. Monte Carlo szimuláció .....	60
7.1.2. Panjer rekurzió .....	66
7.1.2.1. Diszkrétizálás .....	67
7.1.2.2. Panjer rekurzió és a tőke stabilitása .....	68
7.1.3. Fourier transzformáció .....	71
7.2. Single Loss Approximation és tőkeszámítás.....	73
7.2.1. Single Loss Approximation és az extrapolációs probléma .....	78
7.2.2. A tőkebecslés hibája lognormális veszteségeloszlás esetén.....	78
8. Külső adatok elemzése .....	95
8.1. A HUNOR adatbázisról .....	96
8.2. Adatok előkészítése.....	96

8.3. HUNOR statisztikák.....	97
8.4. Extrém értékek elmélete.....	104
8.5. Az eloszlás szélének vizsgálta.....	105
8.6. GPD szimulációs kísérlet.....	114
8.6.1. A tail index eloszlásáról.....	116
8.6.2. A scale paraméter eloszlásáról.....	117
8.6.3. A tőke eloszlásáról.....	117
9. Konklúziók a működési kockázati tőkekövetelmény instabilitásáról.....	124
10. Portfóliók zajérzékenysége.....	128
11. A variancia kockázaterzékenysége.....	132
11.1. A Markowitz probléma.....	132
11.2. A Markowitz probléma geometriai interpretációja.....	133
11.3. Empirikus kovariancia mátrixok.....	135
11.4. Egy érdekes megfigyelés.....	136
11.5. Wishart mátrixok sajátérték eloszlása.....	137
11.6. Egy szűrési eljárás a véletlen mátrixok elméletére alapozva.....	138
11.7. Az optimális portfólióra gyakorolt zaj hatásának mérése.....	140
11.8. Modell építés és eredmények.....	140
11.8.1. 1. modell: A legegyszerűbb kovariancia mátrix.....	140
11.8.2. 2. modell: Az egyfaktoros (piaci) modell.....	142
11.8.3. 3. Modell: Szimmetrikus piac-plusz-szektorok modell.....	144
11.8.4. 4. modell: Antiszimmetrikus piac-plusz-szektorok modell.....	146
12. Alternatív kockázati mértékek.....	152
12.1. Value at Risk és a szabályozói mértékek.....	152
12.2. Abszolút eltérés.....	154
12.3. Expected shortfall.....	158
12.4. Maximális veszteség.....	160
12.5. Az expected shortfall megoldhatósági problémája.....	165
12.6. A súlyok ingadozása és $q_0$ eloszlása.....	173
13. Konklúziók a portfólióoptimalizálásban használt kockázati mértékek instabilitásáról ...	176
14. Saját eredmények összegzése.....	178
15. Befejezés.....	185
Dolgozatban felhasznált saját publikációk jegyzéke.....	187
Summary.....	188
Rövidítésjegyzék.....	194
Hivatkozások.....	196

## *1. Köszönetnyilvánítás*

A dolgozat tudományos eredményei szempontjából alapvető támogatást és iránymutatást kaptam témavezetőim Dr. Gáll József Mihály és Dr. Kondor Imre, valamint Dr. Pafka Szilárd kollégám részéről. Ezúton is szeretném megköszönni az egyes problémákkal (kutatási hipotézisekkel) kapcsolatos kitartó támogatásukat és segítségüket, valamint a közös munkát. Ezúton szeretnék köszönetet mondani a munkahelyi vita bírálóinak, Dr. Kosztopulosz Andreász és Dr. Szüle Borbála egyetemi docenseknek a bírálatban és a munkahelyi vitán megfogalmazott számtalan hasznos megjegyzésükért. Végül, de nem utolsósorban köszönettel tartozom feleségemnek, Papp Katalinnak a hosszú évek türelméért és a számos építő észrevételért.

## 2. Bevezetés

Az elmúlt évtizedek a pénzügyi modellek gyors terjedését hozták a pénzügyi szektorban, ami a bankműködés számtalan területén megnyilvánul. Modelljeink segítségével meghatározzuk számos pénzügyi termék árát, ügyfelet választunk, eldöntjük, hogy adott kockázatvállalási szint mellett melyik befektetés a legjobb az ügyfél számára, vagy éppen kiszámítjuk a bank szükséges tőkeszintjét. A statisztikai modellek építése és alkalmazása igen fontos, hiszen a modellek által kínált többlet információ gyakran versenyelőnyt jelent, vagy az információ birtoklása egyenesen a versenyben maradás feltétele lehet. Ugyanakkor szem előtt kell tartanunk azokat a szempontokat is, melyek a modellek által adott információ értékét csorbítják. Ennek megfelelően tisztában kell lennünk például a modell feltevésekkel, az inputok minőségével és azok releváns voltával (pl. egy sokkuktól mentes pénzügyi idősor biztosan nem fogja visszatükrözni a tényleges kockázatokat), a környezeti változásokkal melyek az inputokra hatnak, és nem utolsósorban a modellezett probléma természetével. Ezek azonban sokszor rejtve maradhatnak a banki modellező előtt, hiszen adott esetben az intézmény az alkalmazott modellt szoftveres támogatással együtt vásárolta, és bízva a szoftver gyártójában, az intézmény már csak inputokkal látja el a modellt és hasznosítja a kapott eredményt. Más esetben bizonyos modellek futótűzként terjednek el egy adott probléma megoldására, és azok piaci alkalmazása lényegesen hamarabb terjed el, mint a modellek az iparág szintű megértése. Az ilyen esetek felgyorsíthatják a modell fogyatékosainak a feltárását annak széleskörű felhasználása révén, egyúttal azonban rendszerkockázatot is jelenthetnek az egész pénzügyi szektorra. A kockázatok széleskörű alul becslése (lásd pl. CDO, devizahitel) rendszerszintű sokkot eredményezhet, melynek egyéni és társadalmi költségei igen magasak lehetnek. Ebből adódóan fontos, hogy tisztában legyünk a modellezett probléma természetével, valamint a modell eredmények felhasználhatóságával. A modellekbe vetett túlzott hit elfedheti a probléma természetéből adódó tényleges kockázatokat, melynek következtében a kockázat viselőjét a naiv tudatlanság állapotában éri a kockázat nem várt mértékének bekövetkezése.

A dolgozat során két olyan pénzügyi problémát mutatunk be, ahol kockázatok modellezése és azok megismerése triviálisnak tűnik, ugyanakkor a modellben rejlő instabilitás következtében az eredmények erősen megkérdőjelezhetőek.

Az dolgozat első részében a működési kockázatokra alkalmazott tőkemodellek stabilitását vizsgálom. Itt fontos hangsúlyozni, hogy nem önmagukban az alkalmazott modellek jelentik az instabilitást, hanem a modellek működési kockázati környezetben történő alkalmazása. Látni fogjuk, hogy definíció szerint a szabályozó által elvárt tőke a működési kockázatok fedezésére olyan elvárást támaszt a pénzügyi intézmények felé, amely magában hordozza az instabilitást. A dolgozat második részében egy portfólióoptimalizálási problémával foglalkozunk. A kockázati mértékek fejlődése töretlen volt az elmúlt 50 évben, mindig favorizálva valamely mértéket a kockázatok számszerűsítésére. Markowitz portfólió modelljében – elsődlegesen marketing okok miatt - a szórás a kockázati mérték. Később a 90-es években a JP Morgan nyilvánosságra hozta az általa alkalmazott Value at Risk (VaR) alapú kockázatkezelési módszertant, ami pillanatok alatt a bázeli szabályozásba is beépült. A VaR kritikái életre hívták a Conditional Value at Risk (CVaR) és spektrális mértékek közé tartozó Expected Shortfall-t (ES). A 2008-ban induló válságot követően a szabályozás egyre nyitottabbá vált az ES széleskörű bevezetésére, mivel az a jelenlegi szabályozó mértéknél (VaR) nagyobb mértékben képes megragadni az extrém veszteségeket. Kockázati mértékre szükség van, hiszen az információt sűrít, egy komplex világot (pl. egy bonyolult bank portfóliót) képez le egyetlen értékke, melynek révén könnyen megismerhetjük annak kockázatait. Ugyanakkor itt is tisztában kell lennünk azzal, hogy egyrészt a kockázati mérték előállításánál nagyfokú egyszerűsítés, információ veszteség történik, másrészt az egyes mértékek a mögöttes világ más és más jellemzőit ragadják meg jól, vagy hangsúlyozzák az elemző számára. Ezt kiegészítve a mögöttes probléma természetével gyakran szembesülhet az elemző, vagy döntéshozó azzal a jelenséggel, hogy miközben a mögöttes világ kockázati profilja változatlan, a kockázatmérés eredménye mégis jelentős eltérést mutat. A dolgozatban vázolt első két probléma pontosan ezzel a jellemvonással bír, míg a harmadik az instabilitás egy másik oldalára világít rá.

Fontos tisztázni, hogy az instabilitás szót egyfajta gyűjtőfogalomként használjuk, semmiképp sem egy egységes, precíz matematikai fogalomként. Ennek megfelelően a dolgozat során a stabilitásnak sincs egzakt definíciója.

Ahogy a disszertáció címe is mutatja mérés és mérték instabilitásáról van szó, hiszen esetenként még az is különbözik, hogy az instabilitás egyáltalán mihez is köthető (méréshez vagy mértékhez?). Az instabilitás fogalmának konkretizálásra az egyes problémák kapcsán kerül sor. Mint azt már felvezettük, a dolgozatban a működési kockázatok és a portfólióoptimalizálás témakörében olyan gyakorlatból merített problémákat taglalunk, melyekkel kapcsolatban a hétköznapi felhasználó meghökken, és azt tapasztalja, hogy a



probléma megoldása esetenként erősen eltér az előzetes várakozásaitól. Tehát valamilyen anomália áll a probléma hátterében. Így az instabilitás a vizsgált esettől függően más és más módon jelenik meg. Működési kockázatok esetén az instabilitás (tőkekövetelmény nagyfokú változékonyságának) elsődleges forrása a szabályozói definíció, ami a modellező szemüvegén nézve gyakran:

- nagyfokú becslési hibában,
- extrémekre való érzékenységben,
- nem megfelelő eloszlás használatában,
- különböző becslési módszerek jelentősen eltérő eredményében

jelenik meg.

Portfólióoptimalizálás esetén az instabilitás egyrészt a portfóliósúlyok nagyfokú változékonyságában, másrészt a portfólió optimalizálhatóságában tükröződik vissza. Azt láthatjuk, hogy a szakaszonként lineáris kockázati mértékek becslési hibája nagyobb, mint a szórásé, és egyben a portfóliósúlyok nagyobb változékonyságával (instabilitásával) járnak. Bizonyos esetekben az instabilitás azt a meglepő ténytet jelenti, hogy az optimalizálás el sem végezhető. Vagy gondolhatunk arra a nem hétköznapi jelenségre, hogy az Expected Shortfall esetében a hiba mértéke nem monoton a konfidenciaszint függvényében. Azaz bizonyos esetekben, ahogy nő az optimalizáláshoz felhasznált mintaelemszám, úgy romlik a becslés.

**A dolgozat főbb kérdéscelvetései a következők.**

- 1. A működési kockázatok tőkekövetelményéhez kapcsolódó extrapolációs probléma ugyanúgy érinti-e a kis és a nagy mintával rendelkező intézményeket?**
- 2. A portfólióválasztási feladatok megoldását mennyiben befolyásolja a választott kockázati mérték? Avagy az egyes mértékek mennyire érzékenyek az optimalizálás során elkövetett becslési hibára?**
- 3. Létezik-e minden esetben optimális portfólió? Azaz a különböző portfólióoptimalizálási feladatoknak létezik-e minden esetben megoldása?**

Hogy a fenti kérdésekre választ adjunk, a dolgozatot az alábbiak szerint strukturáljuk. A dolgozat első részében (3-8 fejezetek) a működési kockázathoz kapcsolódó kérdést vizsgáljuk. Ehhez hosszabb és részletesebb bevezető tartozik, mivel az kevésbé ismert, míg a

portfólióoptimalizálásról szóló második rész nagyjából ismertnek feltételezi az alkalmazott keretrendszert. Ennek oka, hogy a modern portfólió elmélet gyökerei Markowitz [1952] munkásságáig nyúlnak vissza, és azt a legtöbb pénzügyi tankönyv taglalja.

A jelen bevezetést követő fejezetek lebontása a következő.

A 3. fejezetben röviden felvázoljuk a működési kockázatokhoz kapcsolódó szabályozói keretrendszert, és a szabályozó által kínált lehetőségeket, melyekkel egy adott hitelintézet működési kockázati tőkekövetelményét számíthatja. Dolgozatunk fókuszpontjában az ún. AMA (Advanced Measurement Approach) áll, melynek keretében az intézmények saját modellt fejleszthetnek, illetve azt tőkekövetelmény számításra alkalmazhatják.

A 4. fejezetben az AMA keretrendszer kötelező elemeit mutatjuk be, és nagy vonalakban rávilágítunk az azokhoz kapcsolódó dilemmákra. Ezek olyan elemei a belső modelleknek, melyek jelentősen befolyásolják az eredményül kapott tőkekövetelményt, ugyanakkor a szabályozó alapvetően nem ad iránymutatást ezek használatára vonatkozóan, csupán azok kötelező meglétét írja elő.

Ezt követően ismertetjük a működési kockázatok közös jellemzőit, természetét, stilizált tényeit. Ezek közül előzetesen kettőt emelünk ki. Egyrészt a működési kockázati tőkekövetelményt alapvetően a legnagyobb egyedi (évközi) veszteség, vagy veszteségek határozzák meg. Ez már önmagában előrevetít egyfajta instabilitást a tőkekövetelmény számítása során. Másrészt a rendelkezésre álló minta sosem elégséges a megbízható becsléshez (lásd Single Loss Approximation). Ez egy újabb forrása a tőkekövetelmény becsléshez kapcsolódó instabilitásának. Kutatási kérdésfelvetésünk is ehhez kapcsolódik. Ha a minta mérete sosem elégséges, akkor kik vannak szerencsésebb helyzetben, a kisebb, vagy a nagyobb belső mintával rendelkező intézmények, vagy alapvetően nincs különbség? Az olvasó számára első körben abszurdnak tűnhet a kérdésfelvetés is, hogyan merülhet fel egyáltalán, hogy egy nagyobb mintából rosszabb becslés szülessen, mint egy kisebbből. A probléma bemutatását a rákövetkező fejezetekben a pénzügyi szabályozás elvárásaiból származtatjuk.

A fejezet végén a Basel II-es szabályozás igen fontos aspektusára mutatunk rá, miszerint a tényleges kockázatokat az extrém veszteségek jelentik, és a kockázatkezelésnek elsődlegesen erre kell figyelmet fordítania. Ennek megfelelően a hagyományos és modern kockázatmenedzsment filozófiáját is szemléltetjük.

Az 5. fejezet definiálja a működési kockázati tőkekövetelményt, és ahhoz kapcsolódóan röviden bemutatja tőkeszámítási módszerek módszertanát. A Loss Distribution Approach (LDA) egy lehetséges eljárást jelent a tőkekövetelmény meghatározására. Lényege, hogy az

aggregált éves veszteséget az egyedi (éven belüli) veszteségek gyakoriság - és veszteség eloszlás segítségével állítja elő. Az LDA eljárás számos módszerrel kivitelezhető, így a további fejezetekben szót ejtünk a Monte Carlo szimulációról (mc), gyors fourier transzformációról (fft) és Panjer rekurzióról is.

Az LDA-n túli, tőkekövetelmény számításra alkalmas egyéb módszerek kapcsán a számunkra igen fontos Single Loss Approximation (SLA) módszertanát és a közelítésem módszereket említjük.

A 6. fejezetben elsőként egy AMA szerinti működési kockázatkezelési modell főbb építő elemeit vesszük sorra. Szót ejtünk az optimális mintanagyságról, valamint az LDA-t meghatározó főbb gyakoriság és egyedi veszteségeloszlásokról. Egyszerű példákon keresztül rámutatunk arra, hogy tőkekövetelmény meghatározása során a gyakoriság eloszlás alakja nagyjából mellékes, annak csupán várható értéke számít, ugyanakkor az egyedi veszteségeloszlás széle domináns szerepet játszik. Illusztratív példát hozunk a működési kockázatkezelési tőkekövetelmény jelentős instabilitására is. Megmutatjuk, hogy a döntéshozó előtt álló információs halmazból (historikus idősorból) alapvetően nem lehet következtetni a működési kockázatkezelési tőkekövetelmény „igazi” értékére.

Ezt követően egy speciális problémát taglalunk. A belső modellt építő intézmények tipikusan egy adott küszöb fölött gyűjtik a működési kockázati veszteségeket, így a küszöb alatti veszteségek eldobása egyben információ veszteséget is jelent. Egyrészt vizsgáljuk, hogy az eldobott veszteségek materiálisak-e a tőkekövetelmény mértékének szempontjából, másrészt megmutatjuk, hogy a feltételes eloszlás használata a tőkekövetelmény számítás során mennyire instabil helyzetet teremt.

Itt hívjuk fel a mintában lévő azonos értékekre a figyelmet. Ennek megemlítése azért szükséges, mert amint azt a HUNOR adatbázis vizsgálata során látni fogjuk, ez a jelenség az egyedi káreloszlás becslésére szintén torzító hatással van.

A 7. fejezetben térünk vissza az LDA módszer lehetséges megvalósításaira. Vizsgáljuk, hogy a különböző módszerek (Monte Carlo, FFT, Panjer rekurzió) alkalmazása bír-e bármilyen hatással a tőkekövetelmény számításra.

Ezt követően bemutatjuk a Single Loss Approximation és annak továbbfejlesztett módszertanát tőkeszámítás szempontjából, részletezzük az első kutatási problémát. Az SLA módszertanát felhasználva kifejtjük, hogy a kis, illetve nagy mintából történő becslés jóságának dilemmáját az jelenti, hogy két ellentétes hatás áll szemben egymással, melyek meghatározzák a tőke becslési hibáját. A nagyobb mintával (az éven belüli veszteségek számából adódóan) rendelkező intézmények látszólag kedvezőbb helyzetben vannak a becslés

szempontjából, ugyanakkor nekik egy magasabb valószínűséghez tartozó kvantilist kell becsülniük, ami egyidejűleg növekvő bizonytalanságot hordoz. A két egymással ellentétes hatás (növekvő minta és magasabb konfidenciaszint) következtében nem nyilvánvaló, hogy a sok adattal rendelkező hitelintézetek valóban kisebb, vagy lényegesen kisebb hibával becslik-e tőkéjüket. A kutatási kérdés megválaszolásához bizonyos feltételek mellett meghatározzuk a tőkekövetelmény eloszlását, ami a dolgozat egyik önálló analitikus eredményét adja. Látni fogjuk, hogy bizonyos esetekben a tőkekövetelmény (esetünkben VaR) szórása az éves gyakoriság függvényében nem monoton viselkedést mutat (kezdetben csökken, majd nő). Megállapítjuk, hogy nem csak az számít, hogy mennyi adattal rendelkezik egy intézmény, hanem az is, hogy mennyire vastagszélű (heavy tail) a veszteségeneráló folyamat. A kettő összjátéka határozza meg, hogy mennyire pontosan tudja az intézmény megbecsülni a működési kockázati tőkekövetelményét.

Mint azt már korábban említettük két hatás feszül egymásnak. Egyrészt a növekvő gyakoriság kedvez a becslésnek, hiszen nagyobb mintából a veszteség eloszlást kisebb hibával tudjuk becsülni, másrészt viszont a nagyobb gyakoriság egyre távolabbi kvantilis extrapolálását teszi szükségessé. Úgy tűnik, hogy egy kevésbé kockázatos (heavy tail) folyamat esetén a mintanagyság növekedés ellensúlyozni tudja az extrapolációból (egyre magasabb kvantilis becsléséből) adódó bizonytalanságot, addig vastag szélű folyamatok esetében ez már nem teljesül. Ebből adódóan paradox módon a mintanagyság növekedése a tőkekövetelmény becslésének nagyobb bizonytalanságával párosul vastag szélű folyamatok esetében. A további vizsgálatok azonban rávilágítanak, hogy amennyiben a tőkekövetelmény hibáját relatív szórásban mérjük, úgy az monoton csökkenő függvénye lesz a mintaelemszámnak, tehát a nagyobb minta stabilabb becsléshez vezet.

A 8. fejezetben a külső adatokkal foglalkozunk. Egy AMA modellt építő intézmény számára nem csak a belső historikus adatok relevánsak, hanem a külső, iparági adatok is. A fejezetben bemutatunk néhány leíró statisztikát a HUNOR (HUNgarian Operational Risk database) adatbázisról, illetve néhány becslést is elvégzünk GPD (általánosított Pareto) eloszlást feltételezve az egyedi veszteség adatokról. A GPD eloszlás relevanciáját az adja, hogy egy adott küszöbön túl (épp ahol a belső intézményi adatok már ritkák, vagy nincsenek) az eloszlások szélei a GDP eloszláshoz tartanak. Megmutatjuk, hogy a különböző becslési módszerek, még ugyanazon eloszlás feltételezése mellett is lényegesen eltérő tőkekövetelményeket adhatnak, azaz a kockázatkezelésnek igen óvatosnak kell lennie az eredmények tekintetében. Mind a HUNOR, mind mesterségesen generált adatok fölött vizsgáljuk a népszerű MLE, és stabil becslést adó PWM becslések közötti különbségeket, és

számos megállapítást teszünk az általánosított Pareto eloszlásból becsült paraméterekre és az azokkal számított tőkekövetelményekre.

Összegzésképpen vázoljuk a konklúziókat (9. fejezet) a működési kockázatkezelési modellek építésével, és az azokból adódó tőkekövetelménnyel kapcsolatban. Látni fogjuk, hogy AMA modell építésekor az intézménynek igen körültekintően kell eljárnia, hiszen egy olyan szabályozói környezetben kell a tőkét meghatároznia, ahol a tőkeérték messze a megfigyelhető és visszamérhető tartományon túl helyezkedik el.

A 10. fejezet már a portfóliók optimalizálása során elkövetett becslési hibákkal foglalkozik. Itt túllépünk a korábban alkalmazott kockázati mértéken (Value at Risk, VaR), sőt éppen ezzel nem foglalkozunk annak nem konvex jellege miatt. Mivel a portfólióoptimalizálás kérdésköre széleskörben ismert és tárgyalt, így itt az általános bevezetést mellőzve, rövid és mélyen specifikus bevezetés után vezetjük fel a második kutatási kérdésfelvetés vizsgálatát, azaz hogy a portfólióválasztási feladatok megoldását mennyiben befolyásolja a választott kockázati mérték?

A 11. fejezet a klasszikus kockázati mérték, a variancia zajérzékenységét vizsgálja. Szimulációs megközelítés alapján vizsgáljuk a variancia zajra való érzékenységét (Pafka et. al. [2002]) alapján. Pafka et. al. [2002] szimulációs környezetben bevezette a becsült kovariancia mátrixú és a becslési hibától mentes portfóliók szórásának hányadosát, és használta a zaj mértékének megállapítására. Megmutatta, hogy a zaj  $N/T$  (portfólió-elemszám és a mintahossz hányadosának) függvénye. Reprodukáljuk az általa felépített variancia-kovariancia mátrix modellt, ami összefogja a valódi piacon megfigyelt empirikus kovariancia mátrixok lényeges tulajdonságait.

A 12. fejezetben térünk ki az alternatív kockázati mértékekre. A második kutatási kérdésfelvetés mentén vizsgáljuk az abszolút eltérést, Expected Shortfallt és ennek pesszimista verzióját, a Maximális Veszteséget (Maximal Loss, ML). Ezen kockázati mértékeknek az elterjedt szabályozói mértékekkel (itt nem a VaR-ra, hanem a mesterségesen szabályozói súlyokon alapuló mértékekre gondolunk) közös az a tulajdonsága, hogy részben lineárisak, vagy, máshogy megfogalmazva iso-kockázati felszínük poliéder a portfóliósúlyok függvényében. Azt állítjuk, hogy ez a tulajdonság a magyarázata annak, hogy ezen kockázati mértékek nagyobb mértékben érzékenyek a zajra a szóráshoz képest. Egy másik döntő faktor, hogy a VaR, ES és ML egy adott küszöb alatt az összes adatot elveti, azaz csak az adatok egy részének jut szerep a becslés során. Megjegyezzük, hogy a variancia és Absolute Deviation mértékekkel kapcsolatos eredmények csupán reprodukciói más szerzők (elsődlegesen Dr. Pafka Szilárd) munkáinak, illetve az alkalmazott szimulációs keretrendszer is átvételre került.

A fejezet végén áttérünk a harmadik kutatási kérdésfelvetés kapcsán arra, hogy az Expected Shortfall fölötti optimum egy valószínűségi kérdés, amely a mintától függ. Mindez azt is jelenti, hogy a portfólióoptimalizálási feladatnak nem minden esetben létezik megoldása. A Maximális Veszteség esetében meghatározzuk az optimalizálhatóság valószínűségét N/T függvényében.

Konklúzióval zárjuk a portfólióoptimalizáláshoz kapcsolódó fejezeteket. Látni fogjuk, hogy a vizsgált kockázati mértékek különbözőképpen érzékenyek a zajra. Azt tapasztaljuk, hogy az Expected Shortfall és annak speciális esete a Maximális Veszteség esetében a legnagyobb a becslési hiba. Az Expected Shortfall és ahhoz kapcsolódóan a Maximális Veszteség esetében az optimum létezését semmi sem garantálja. A megoldás létezése kikényszeríthető további korlátozó feltételek bevezetésével, de ez már az eredetitől eltérő lineáris programozási feladatot jelent.

Legvégül összegezzük a kutatási kérdésekre adott válaszokat, és a dolgozat saját eredményeit. A dolgozatot a köszönetnyilvánítás, összegzés, és hivatkozások fejezetei zárják.

A dolgozat során számos problémával találkozik az olvasó. Hangsúlyozzuk, hogy nem kínálunk megoldást a problémák orvoslására (hogyan lehetne az adott kontextusban fennálló becslési hibát csökkenteni), elsődleges célunk a probléma megértése, leírása és ezeken keresztül a kutatási kérdésfelvetések körbejárása. Az önálló eredmények nem válnak el élesen a szöveg egyéb részeitől, ugyanakkor a fejezetek, alfejezet címében ezt mindig jelezzük. A dolgozat több ponton is számos olyan eredményt tartalmaz, ami közös eredményt jelent más társszerzőkkel (Dr. Gáll József, Dr. Kondor Imre<sup>1</sup>, Dr. Pafka Szilárd, Hajnal Béla, Kállai Zoltán, Povilaitis Katalin), így itt is ahhoz tartjuk magunkat, hogy az önálló eredményt a címben jelöljük. A dolgozat során végig problémát jelent a tizedespont és tizedesvessző használata, mivel az alkalmazott szoftverek tizedespontot használnak, így az azokból származó ábrákon is tizedespont jelenik meg. Ugyanakkor a szövegben a magyar helyesírási szabályoknak megfelelően tizedesvesszőt használunk, míg a pont gyakran az ezresek szeparálására szolgál. Problémát jelenthet még az olvasó számára, hogy a pénzügyi intézmények szabályozása egy folyamatos tevékenység, így a szabályozói, törvényi hivatkozások a dolgozat, vagy a dolgozat alapját képező publikáció megírása óta változhattak. Ennek következtében előfordulhat, hogy a hivatkozott szabályozói anyag már nem érhető el

---

<sup>1</sup> Témavezető a portfólióoptimalizálás témakörben.

web-es hivatkozás esetén, mivel az hatályát veszítette egy módosítás, vagy magasabb rendű szabályozás bevezetése következtében.

### 3. Működési kockázatok tőkeszámításának lehetőségei

A működési kockázatok kezeléséről és tőkekövetelményéről szóló uniós szabályozás (Capital Requirement Directive, CRD<sup>2</sup>) újabb kihívás elé állította a hitelintézeteket<sup>3</sup> (Validációs Kézikönyv, Pénzügyi Szervezetek Állami Felügyelete (2006), 200/2007 (VII. 30.) Korm. r.). Míg a szabályozó eredetileg kisebb tőkekövetelménnyel kívánta jutalmazni a kockázataikat mélyebben feltérképező, jobban megértő és szofisztikáltabb tőkeképzési módszereket alkalmazó intézményeket, ma már a 2008-ban kezdődő válság következményeként a szabályozó azt gondolja, hogy rendszerkockázatilag fontos intézményeknek kockázaterzékeny módszertan keretei között a működési kockázataikat kezelni és tőkét képezni. Az ilyen modellek implementálása számos nehézséggel jár többek között az idősor nem stacionárius jellege és a nem elégséges mintaelemszám miatt (Embrechts et. al. [2003]). A dolgozat első részében a működési kockázatok természetét, a kockázatokra képzett tőke stabilitását vizsgálom AMA (Advanced Measurement Approach) környezetben számos szimulációs példán, a HUNOR bankközi adatbázis adatain és egy leegyszerűsített modellen keresztül. Az AMA mellett a szabályozó (Pénzügyi Szervezetek Állami Felügyelete [2006]) még másik 3 egyszerűbb tőkeképzési lehetőséget kínál a hitelintézeteknek, az alapmutató módszert (Basic Indicator Approach, BIA), a sztenderdizált módszert (Standardised Approach, TSA) és az alternatív sztenderdizált módszert (Alternative Standardised Approach, ASA). Míg a BIA, TSA és az ASA egyszerű formulákon keresztül adja meg a képezendő tőke nagyságát, addig az AMA keretein belül az intézmény szofisztikált belső modellt építhet a kockázatok valódi jellegének feltérképezésén alapulva.

A fentiek következménye a BIA, TSA és az ASA csekély kockázaterzékenysége, illetve hogy bevezetésünk kezdeti éveiben az AMA-hoz viszonyítva tipikusan magasabb tőkekövetelményt is jelentettek<sup>4</sup>. Ugyanakkor jegyezzük meg, hogy egyes nagy alaptőkéjű hitelintézeteknél néhol előfordultak ezzel ellentétes tapasztalatok is (ld. Committee of

---

<sup>2</sup> CRD alatt az Európai Unió 2006/48/EC számú és 2006/49/EC számú direktíváit értjük.

<sup>3</sup> Hpt. szerinti hitelintézetek, és hitelintézettel egyenértékű prudenciális szabályozásnak megfelelő pénzügyi vállalkozások.

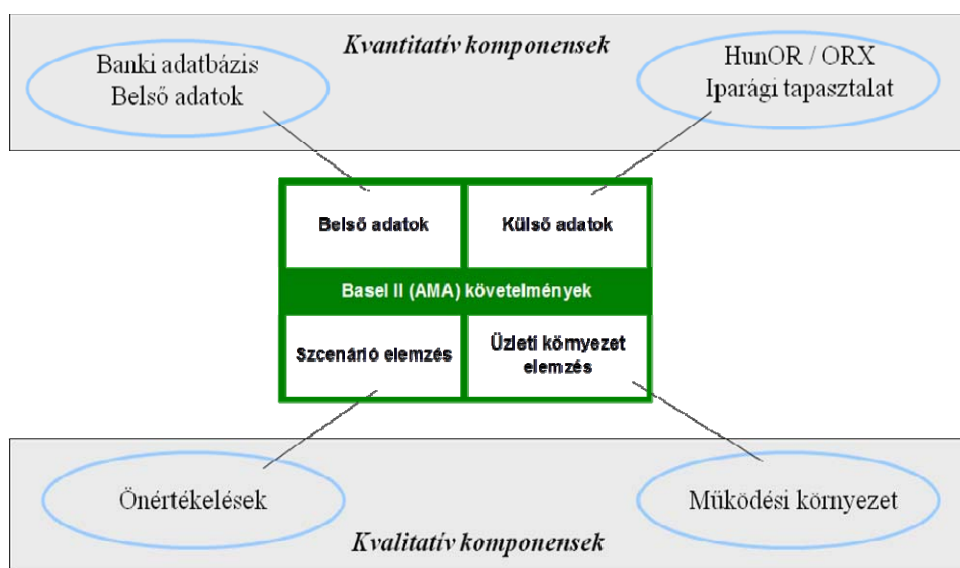
<sup>4</sup> A szabályozás egyik jellemvonása, hogy a szofisztikáltabb módszereket alkalmazó intézményeket kevésbé büntetik magas tőkekövetelménnyel. E mögött az a logika húzódik meg, hogy minél komplexebb módszert választunk, annál pontosabb képet kapunk az intézmény működési kockázatairól. Ugyanakkor a 2008-ban kezdődő pénzügyi válság árnyalta némelyest ezt a képet, hiszen a válság alatt a hitelintézetek működési veszteségei és az abból számolt tőke tipikusan nem csökkent, míg a BIA, TSA, ASA alapjául szolgáló Gross Income igen, ami az ezen módszertanokat használók számára kedvezőbb tőkehelyzetet teremtett.



European Banking Supervisors [2006]). A 2008-ban kezdődő pénzügyi válság tovább árnyalta ezt a képet, hiszen a válság alatt a hitelintézetek működési veszteségei, és az abból számolt tőke tipikusan nem csökkent, míg a BIA, TSA, ASA alapjául szolgáló Gross Income gyakran igen, ami az ezen módszertanokat használók számára kedvezőbb tőkehelyzetet teremtett.

Ahhoz, hogy a hitelintézetek a fejlett módszertant alkalmazhassák számos előírásnak kell eleget tenniük (Basel Committee on Banking Supervision [2003], 200/2007 (VII. 30.) Korm. r.). Ezeket nem soroljuk fel tételesen csupán azokat, amelyek a modellépítés szempontjából relevánsak.

1. A belső modellnek mind a várható mind a nem várt veszteségeket meg kell ragadnia.
2. A kis valószínűséggel bekövetkező, ugyanakkor potenciálisan nagy hatást okozó eseményekre is fedezetet kell nyújtania a tőkének, 99,9 százalékos valószínűséggel, 1 éves időtávon.
3. A belső modellnek figyelembe kell vennie
  - az intézmény saját, belső adatait,
  - a külső adatokat,
  - az üzleti környezetet tükröző tényezőket,
  - a forgatókönyv elemzést (scenario analysis).



1. ábra: A belső modell komponensei.

4. Az intézményeknek legalább 5 éves idősort kell figyelembe venniük a tőkekövetelmény számításához,<sup>5</sup> továbbá a hitelintézet belső veszteségadatainak átfogónak<sup>6</sup> kell lenniük.
5. A hitelintézeteknek a veszteség adatokat üzletágaknak és veszteség kategóriáknak kell megfeleltetnie.
6. A belső adatok gyűjtésének alsó határértékét az intézménynek meg kell határoznia.
7. A hitelintézet modelljébe beépíthet és így alkalmazhat a veszteségadatok között korrelációs feltételezéseket<sup>7</sup>.

A Basel II-es keretrendszer<sup>8</sup>, melynek célja a nemzetközi bankrendszer stabilitásának az előmozdítása számos gyakorlati problémát vet fel, melyeket az intézmény belső, másrészt a külső iparági adataihoz köthetünk (Basel Committee on Banking Supervision [2011]).

#### *4. Basel II keretrendszer - AMA*

##### *4.1. Belső adatok*

Egy intézmény működési tőkekövetelményének meghatározása céljából igen fontos szerepet játszanak annak belső, szervezetből származó adatai. Ezek az adatok, működési veszteségek a szervezetben szétszórtan vannak jelen, így a működési kockázatkezelés fontos feladata az adatok gyűjtésének megszervezése. A piaci kockázatkezeléssel ellentétben az elemzéshez / tőkeszámításhoz szükséges adatok nem vásárolhatók meg valamely információ szolgáltatótól (Reuters, Bloomberg), hiszen mások által nem ismert, a szervezet saját belső adatairól, veszteségeiről van szó. Az adatok gyűjtése, rendszerezése igénybe veszi az egész szervezetet, és értelemszerűen adódik, hogy nem éri meg minden apró veszteséget feltárni és adatbázisban rögzíteni. Ennek következménye, hogy a szervezet meghatározza azt az alsó küszöbértéket,

---

<sup>5</sup> A fejlett mérési módszer bevezetésekor elegendő 3 évnyi adattal rendelkeznie a hitelintézetnek.

<sup>6</sup> Átfogónak kell a belső veszteségadatoknak lenniük abban az értelemben, hogy azoknak meg kell ragadniuk a vonatkozó alrendszerek és földrajzi régiók összes főbb tevékenységét és a kitétségeit. A hitelintézeteknek bizonyítaniuk kell, hogy a kizárt tevékenységek vagy kitétségek sem egyéneként sem együttesen nem befolyásolják lényegesen az átfogó kockázati becsléseket.

<sup>7</sup> Amennyiben tudja igazolni, hogy a korreláció mérésére alkalmazott módszerei megbízhatóak (azaz mennyiségi és minőségi módszerekkel alá vannak támasztva), továbbá azt, hogy figyelembe veszi a korrelációs becslések ismert hiányosságaiából adódó hibákat (200/2007 (VII. 30.) Korm. r.).

<sup>8</sup> A Basel II után mélyebben érdeklődő olvasónak javasoljuk Ong [2006] könyvét.

ami fölött veszteségeit gyűjti. A későbbiekben ezzel kapcsolatban még visszatérünk arra a kérdésre, hogy az adatgyűjtési küszöb megválasztása hogyan hat a működési kockázati tőkére.

#### *4.1.1. Dátum választás*

Az adatok gyűjtéséhez kapcsolódó problémát jelent a modellezési dátum kiválasztása. A szervezet számos időpontot - esemény bekövetkezésének dátuma, esemény észlelésének dátuma, eseményhez tartozó veszteség könyvelésének a dátuma, esemény adatbázisban történő rögzítésének a dátuma feljegyezhet egy működési kockázati eseménnyel kapcsolatban, melyek használata különböző tőkeszámokat eredményezhet, tehát a szervezetnek körültekintően kell választania közülük. Egy működési kockázati esemény bekövetkezése sokszor nem ismert, így adott esetben az sem eldönthető, hogy az eseményhez tartozó veszteségnek bele kell-e tartoznia abba az idősorba, amiből a tőke meghatározása történik. Az észlelés dátuma egy újabb lehetőség a működési kockázati esemény hatásának modellezésére, hiszen akkor vesszük figyelembe az eseményt, amikor már tudjuk, hogy viselnünk kell annak következményeit (legtöbbször a bekövetkezett veszteséget detektáljuk), ugyanakkor események melyek a (rég) múltban következtek be, bekövetkezésük után sokkal később is okozhatnak gazdasági hatást. Ez utóbbi esetén olyan események kerülnek be az idősorba és a tőkeképzésbe, melyek adott esetben már nem is relevánsak az intézmény szempontjából (pl. az üzletág már megszűnt, vagy azt a folyamatot már számos kontroll funkció erősíti, ahol az esemény bekövetkezett).

A könyvelés és rögzítés dátuma mindössze technikai dátumok, így egy kockázat érzékeny tőkekövetelmény megállapítására nem javasoltak.

#### *4.1.2. Időtáv választás*

A szabályozó 5 éves idősort javasol a működési tőke meghatározásához, ugyanakkor az idősor hosszának meghatározása során azt szükséges mérlegelni és optimalizálni, hogy az intézménynek legyen elégséges adata a kockázatok becsléséhez, tehát egyfelől legyen minél hosszabb az idősor, másrészt nem szabad arról sem megfeledkeznie, hogy egy állandóan változó környezetben működik, tehát a valós kockázati profilját az időben legfrissebb események jellemzik leginkább. Ezt a képet árnyalja, hogy lehetnek olyan típusú események

(pl. bírósági ügyek), ahol a veszteségek és a kapcsolódó megtérülések időben távol vannak egymástól, így egy rövidebb – akár 5 éves időtáv választása értelmét veszti.

#### *4.1.3. Modellezett mennyiség megválasztása*

A szabályozó ugyan előírja, hogy figyelembe kell venni a megtérüléseket a tőke számítása során, ugyanakkor annak módját csak részben szabja meg. A szereplők egy része a megtérülés (kivéve biztosítások) utáni, mások a megtérülés előtti összeget tekinti a modellezés alapjának.

#### *4.1.4. Csoportos veszteségek*

A hitelintézet köteles rendelkezni az egymással összefüggő, azonos okra visszavezethető veszteségekről. Ennek legegyszerűbb formáját az ugyanazon működési kockázati esemény által generált veszteségek jelentik, míg ennél nehezebben megragadhatók az ugyanazon ok miatt bekövetkező események veszteségei. Ez utóbbiak közé tartoznak azon összetett veszteségek, melyek sorozatban következnek egymás után ugyanazon esemény miatt, illetve az ún. makro események, melyeket a közös kiváltó ok kapcsol össze. Csoportos veszteségek esetén dönteni kell a kapcsolódó veszteségek aggregálásáról vagy egyedi értéként kezeléséről a modellben.

#### *4.1.5. Adatgyűjtési küszöb*

A belső adatokra vonatkozóan az intézménynek meg kell határozni egy minimális veszteségküszöböt, amely fölött veszteségadatot gyűjt. A 200/2007-es kormányrendelet alapján az adatbázisból csak azok a veszteségadatok hagyhatók ki, melyek bizonyíthatóan nem befolyásolják jelentősen sem egyedileg, sem pedig összességében a teljes kockázatot.

### *4.2. Külső adatok*

Az intézmény belső adatai között extrémek – igen nagy veszteségek alig, vagy legalábbis modellépítés szempontjából tipikusan elégtelen mennyiségben fordulnak elő. Ugyanakkor

kockázat és tőkekövetelmény szempontjából éppen ezek az események a relevánsak (lásd Aue et. al. [2006], Böcker et. al. [2005], Embrechts et. al. [2003]). Az intézmény külső adatai két forrásból származhatnak:

- publikus adatok,
- konzorcium adatok.

Mindkét adatforrás rendelkezik előnyökkel és hátrányokkal, a kettő közötti választást elsődlegesen a felhasználás célja dönti el. A publikus adatok gyakran részletes leírással rendelkeznek az eseményt illetően, ugyanakkor közzétételük meglehetősen eseti. Ebből adódóan hasznosak lehetnek a kockázati önértékelések és scenárió elemzések során, ugyanakkor eseti jellege miatt nem képezheti stabil inputját a tőkeszámításnak. A konzorcium adatok rendszeres és folyamatos rendelkezésre állását a tagok azon vállalása biztosítja, miszerint azok meghatározott időközönként közzéteszik működési kockázati veszteségeiket a konzorcium többi tagjának számára. Ezen adatforrás hátrányát éppen az jelenti, hogy a tagok anonimak szeretnének maradni a veszteségek tekintetében, így a kockázati eseményekről tipikusan csak az adatkonzorcium által előírt minimális információkat adják meg. A felhasználók így sokszor kénytelenek megbízni a többi tag adatgyűjtési, rögzítési, kategorizálási folyamatában, hiszen az esemény / veszteség leírásából sokszor már nem lehet rekonstruálni a teljes működési kockázati eseményt, így a felhasználó csak annak helyességében bízhat.

A külső adatok bevonása a modellbe nem lehet minden megfontolás nélküli. Az adatkonzorcium tartalmazhat olyan adatokat, melyek egyáltalán nem relevánsak az azt felhasználó számára (pl. nagyon eltérő működési környezet; üzletági veszteségek, ahol az adott üzletággal az intézmény nem rendelkezik), így azokat mellőzni praktikus a tőkeszámítás során. A relevánsnak tűnő adatok bevonása pedig azt eredményezheti, hogy a tőkekövetelmény az iparági tőkekövetelmény irányába fog torzítani. A tőke becslésének külső adatok miatti torzítását annak stabilitása kompenzálhatja külső adatok figyelembe vétele esetén. A kérdés tehát egy intézmény számára, hogy az mennyire hasonlít, vagy különbözik az iparágtól, hiszen ennek függvényében fogja használni (súlyozni) a külső adatokat. Számos szerző (ld pl. Dahlen et. al. [2008]) javasolja a külső adatok skálázását (intézményre szabását), ami egy az iparági tőkekövetelmény felé kevésbé torzító tőkeértékhez vezethet.

Az bankszektor tipikus megoldásait és tapasztalatait fenti problémák kapcsán (Basel Committee on Banking Supervision [2011] foglalja össze.

### 4.3. *Stilizált tények*

A Basel II-es javaslatok következtében számos intézmény kezdte el vizsgálni a működési kockázatok közös jellemzőit, és annak természetét stilizált tényeknek nevezzük (Embrechts et. al. [2003], Cope et. al. [2009]). Ezek közül a legfontosabbak:

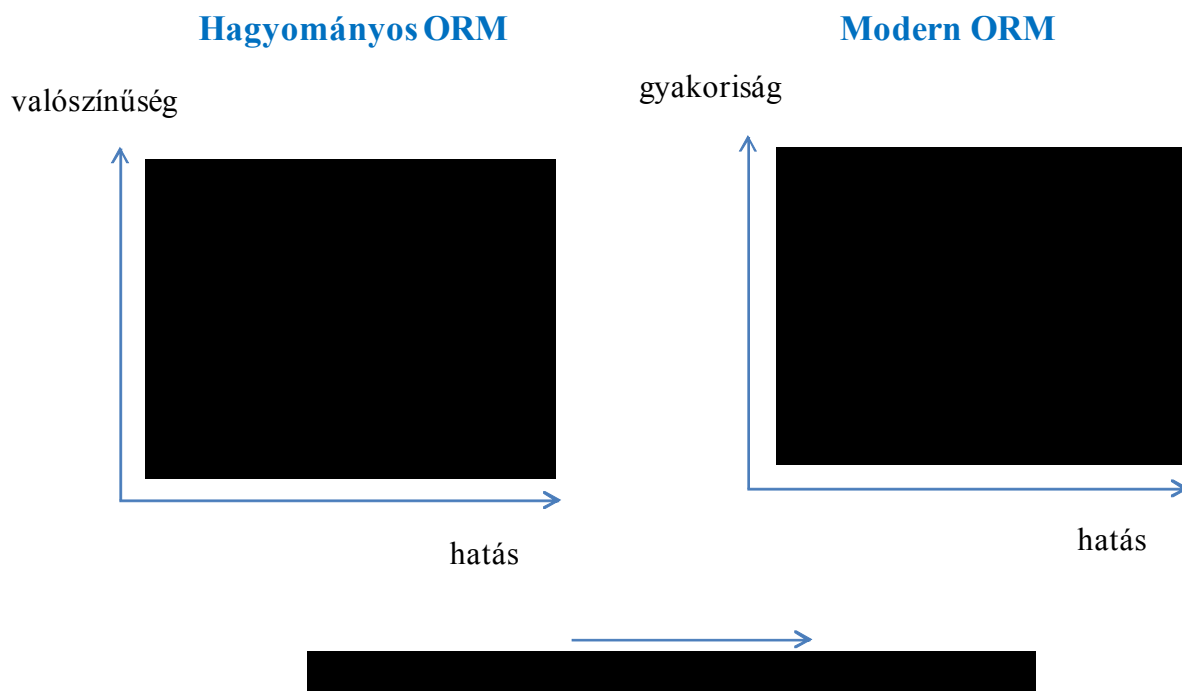
1. A működési veszteségek között extrém értékeket találunk, a veszteségek eloszlása vastag szélű. Ez azt jelenti, hogy a legnagyobb veszteségek nagyságrendje még mérsékelt (néhány száz) mintanagyság mellett is több ezer- tízezerszerese lehet a minta mediánjának. Ezen állítás további következményei, hogy
  - a. az egy adott időszakra aggregált veszteségeket a minta maximális eleme, esetleg elemei határozzák meg;
  - b. a paraméterbecslések instabilak. Egyetlen extrém veszteség nagymértékben képes a minta várható értékét, szórását és egyéb jellemzőit módosítani.
  - c. a heterogén folyamatok nagy veszteségei dominálják a tőkét. Amennyiben két különböző kockázati osztály veszteségeit egy mintaként kezeljük, akkor az extrémeket tartalmazó osztály kockázata fogja jellemezni az új, összevont osztályt is.
2. A veszteségek bekövetkezése semmilyen időbeli szabályosságot nem mutat.
3. Az események gyakorisága nem stacionárius folyamatot követ.
4. A minta mérete mindig elégtelen a tőke meghatározásához. Ennek kifejtése később a Single Loss Approximation keretei között történik.

Mindezek az állítások komoly hatással bírnak az intézmény tőkebecslésére, és amint azt látni fogjuk, a működési kockázati tőke statisztikai alapon történő megközelítése nem kis kihívás elé állítja a bankrendszert.

#### 4.4. Kockázati filozófia

A Basel II által elindított folyamatok, köztük a működési kockázatok megismerése, feltérkepezése, elemzése nagymértékben hozzájárult a kockázati filozófia, gondolkodásmód alakításához. A működési kockázatok kezelése nem új keletű egy intézmény számára, hiszen a működés javítása, folyamatok kontrollja, belső ellenőrzés magával a szervezettel egyidős. Ugyanakkor ezek kockázatszempontrú megközelítése, és az arra a kérdésre keresett válasza, hogy mi is határozza meg a működési kockázati tőkét (ami tulajdonképpen egy kockázati mérték), új logikát alakított ki. Az így kialakult új logikát nevezzük el modern, míg az előzőt hagyományos működési kockázat menedzsmentnek (ORM) Society of Actuaries [2010] alapján. Nem célunk a kettő közötti különbség teljes körű felsorolása, itt csak a dolgot szempontrúából lényeges vonását mutatjuk be.

Mindkét megközelítés egyetért abban, hogy a kockázat mérése során figyelembe kell venni egy esemény bekövetkezési valószínűségét / gyakoriságát és hatását, ugyanakkor a hagyományos megközelítés (lásd pl. COSO) kockázati mátrixa azon eseményeket tekintette veszélyesnek a szervezet szempontjából, ahol mind az esemény bekövetkezésének valószínűsége, mind pedig hatása magas (Patchin et. al. [2012]). A modern megközelítés alapján ilyen események azonban nem léteznek (fiktív kockázat, lásd Samad-Khan [2005] és Samad-Khan [2008]), hiszen túl nagy veszteségeket nem szenvedhet el tetszőleges gyakorisággal a szervezet, máskülönben csődbe megy. A modern megközelítésben a hangsúly elsődlegesen azon eseményekre tevődik, melyek bekövetkezési gyakorisága ugyan kicsi, de hatása a szervezetre magas (extrémek). A következő ábra mátrixai a hagyományos és modern kockázatmenedzsment filozófiáját tükrözik vissza.



2. ábra: Hagyományos és modern kockázatmenedzsment kockázati mátrixai.

## 5. A tőkekövetelmény-számítás filozófiája

### 5.1. A tőkekövetelmény LDA esetén

A pénzügyintézetben az operációs kockázat számításához először részekre kell osztani a operációs kockázathoz tartozó tevékenységek és folyamatok egészét. Ezen részeket elsősorban a szabályozó által ajánlott *veszteségkategóriák* szerint -7 kategóriát szokás kialakítani<sup>9</sup> - és *üzletágak* szerint -8 üzletágot szokás elkülöníteni<sup>10</sup> - érdemes kialakítani. Ezen két szempont szerint tehát létrejön egy 7x8-as méretű mátrix, nevezzük ezt *operációs kockázatok mátrixának*, mely a feltüntetett szempontok szerint egy csoportosítását adja a működési kockázati eseményeknek. Ennélfogva akár 56 különböző csoportot is kialakíthatunk, de a valóságban nem feltétlenül tanácsos egy ilyen részletes bontás, mert annak számos hátránya van, amelyeket a későbbiekben tárgyalunk. Ezért azt javasoljuk Gáll et. al. [2007] alapján, hogy célszerű összevonni egyes elemeit ennek az 56 elemű mátrixnak, s

<sup>9</sup> (a) belső csalás, (b) külső csalás, (c) munkáltatói gyakorlat és munkabiztonság, (d) ügyfél, üzleti gyakorlat, marketing és termékpolitika, (e) tárgyi eszközökben bekövetkező károk, (f) tevékenységbeli zavar vagy rendszerhiba, (g) végrehajtás, teljesítés és folyamatkezelés.

<sup>10</sup> (a) vállalati pénzügyek, (b) kereskedés és értékesítés, (c) lakossági közvetítői tevékenység, (d) kereskedelmi banki tevékenység, (e) lakossági banki tevékenység, (f) fizetési és elszámolási tevékenység, (g) a pénzügyi szolgáltatás közvetítése (ügynöki) tevékenység, (h) vagyionkezelési tevékenység.



így létrehozni néhány nagyobb csoportot. A továbbiakban ezeket a csoportokat az operációs kockázat *osztályainak* fogjuk nevezni. Ezek, ahogy az 56 csoportja a működési kockázatok mátrixának, valóban egy osztályozását adják a lehetséges eseményeknek (veszteségeknek): azaz minden operációs kockázat szempontjából felmerülő esemény beletartozik valamelyik osztályba, ugyanakkor átfedések nincsenek az osztályok között. Nem foglalkozunk viszont ebben az írásban azzal a problémával, hogy egy esemény több osztályt is érinthet (pl. több üzletág együttes hibájának a következménye a veszteség), és ennek következtében az intézmény a veszteséget megosztaná az érintett osztályok között. A továbbiakban jelölje  $M$  a létrehozott osztályok számát. A dolgozatban még érintjük, hogy milyen elvek szerint érdemes az operációs kockázat mátrixának elemeinél az összevonásokat elvégezve kialakítani az osztályokat.

A továbbiakban az egyszerűség kedvéért feltételezzük, hogy adott egy operációs kockázati osztály, s az operációs kockázatot (tőkekövetelményt) erre az osztályra kívánjuk meghatározni. Továbbá feltételezzük, hogy rögzített egy időintervallum (tipikusan 1 év), amelyre meghatározzuk a tőkekövetelményt.

Az LDA módszertana szerint a következőt feltételezzük az operációs veszteségekről. Jelölje  $X_i$  a vizsgált időszakban az adott kockázati osztályban bekövetkező  $i$ -edik eseményhez tartozó (egyedi) *veszteség* értékét (ahol  $i$  pozitív egész). Az egyszerűség kedvéért a későbbiekben azonban azt feltételezzük, hogy egy eseményhez csak egy veszteség tartozik. Ez nem jelent megszorítást, csupán lehetővé teszi számunkra, hogy az esemény és veszteség szavakat szinonimaként használjuk. Ezeket egyedi veszteségeknek is fogjuk a későbbiekben nevezni, hangsúlyozva a teljes veszteségtől való különbséget. Ekkor  $X_i$  egy nemnegatív értékű valószínűségi változó. Feltételezzük, hogy  $X_1, X_2, X_3, \dots$  függetlenek és azonos eloszlásúak. Ezek nem túlzottan szűkítő feltételezések, hiszen ezek értelmében a vizsgált időszakban bekövetkező veszteségek egymástól függetlenek és azok azonos eloszlását azért indokolt feltételeznünk, mert azok ugyanazon rögzített veszteségkategória és üzletág veszteségei, tehát azonos típusúak. Az adatgyűjtéshez és a modellezéshez fontos a veszteségek (események) pontos definiálása és besorolása. Ez különösen lényeges a több veszteségkategóriát vagy üzletágot is érintő eseményeknél, ahol a belső szabályozásnak egyben hatása lehet a tőkekövetelményre, az aggregációs kérdésekre, melyekre később még

utalunk. Jegyezzük meg azt is, hogy egyes események akár negatív veszteséggel is járhatnak, melynek figyelembe vételére jelen írásban nem térünk ki<sup>11</sup>.

Jelölje továbbá  $\eta$  az adott időszakban az adott kockázati osztályban bekövetkező veszteségek számát. Ennélfogva  $\eta$  is egy valószínűségi változó –hiszen nem ismerjük előre a veszteségek számát-, mely nemnegatív egész értékeket vehet fel. A továbbiakban  $\eta$ -t egyszerűen *gyakoriságnak* fogjuk nevezni,  $\eta$  eloszlását pedig *gyakoriságeloszlásnak*. Feltesszük, hogy az  $X_i$  változók az  $\eta$  változótól is függetlenek. Jelölje végül  $S$  az adott időszakban az adott kockázati osztályban bekövetkezett *összes (vagy teljes) veszteség* értékét. Nyilvánvalóan

$$S = \sum_{i=1}^{\eta} X_i . \quad (1)$$

Ahogy hangsúlyoztuk, a fentiekben leírt modell nem a pénzintézet teljes operációs kockázatára, hanem csak egy rögzített (veszteségkategóriák és üzletágak alapján kialakított) osztály kockázatára és ahhoz tartozó tőkekövetelményének meghatározására vonatkozik.

A teljes veszteség itt ismertetett modellje a valószínűségszámításban és statisztikában jártas olvasók számára egy közismert modellt és megközelítést mutat. A későbbiekben tárgyaljuk, hogy milyen eloszláscsaládokat szokás használni a veszteségek és a gyakoriság esetén. Ha specifikusan  $\eta$  eloszlása Poisson, akkor  $S$  eloszlása nem más, mint az ún. *összetett Poisson eloszlás*, melyet a pénzügyi és a biztosítási matematikában is számos helyen alkalmaznak. Általánosan pedig  $S$  eloszlását a továbbiakban *összetett eloszlásnak* fogjuk az egyszerűség kedvéért nevezni. Hangsúlyozni érdemes a függetlenség feltételét a definícióban, amely nem csak az egyedi veszteségekre vonatkozik, hanem azoknak a gyakoriságot leíró változótól való viszonyára is.

---

<sup>11</sup> Ám megjegyezzük, hogy ilyen módon is bővíthetjük a jelen írásban leírt modelleket (például alkalmas feltevésekkel a feltételes veszteségeloszlásokról). Itt említhető az a rokon probléma is, hogy a szabályzó a belső adatokra vonatkozóan veszteségküszöb használatát is lehetővé teszi. Továbbá fontos, hogy az adatbázisból csak azok a veszteségadatok hagyhatók ki, amelyek bizonyíthatóan nem befolyásolják jelentősen sem egyedileg, sem pedig összességében a teljes kockázatot (Validációs Kézikönyv, Pénzügyi Szervezetek Állami Felügyelete (2006)).

A szakirodalomban (Emrechts et. al. [2003], Panjer [2006]) számos tulajdonság, elméleti eredmény ismert az összetett (Poisson) eloszlásokról. Nem célunk ezen írásban az elméleti eredmények összefoglalása, ám a későbbi részekben megemlítjük a szóban forgó eloszlások néhány számunkra szükséges fontosabb tulajdonságát.

A továbbiakban egy  $Y$  valószínűségi változó eloszlásfüggvényét  $F_Y$  fogja jelölni, azaz  $F_S(x) = P(S < x)$ , ahol  $x \in R$ , továbbá  $R$  a valós számok halmaza.

Az összes veszteséghez tartozó *tőkekövetelmény hozzájárulás* alatt annak egy adott biztonsági szinthez tartozó *Value at Risk (VaR, Kockázatotott érték)* értékét tekintjük. Ezt a tőkekövetelmény hozzájárulást a későbbiekben az egyszerűség kedvéért tőkekövetelménynek fogjuk nevezni.

Legyen  $0 < \alpha < 1$  és tekintsük az  $1 - \alpha$  biztonsági szinthez tartozó Value at Risk (VaR) értéket, amely megadja az LDA alapján az adott kockázati osztályhoz tartozó tőkekövetelményt. Itt eltekintünk attól, hogy a szabályozó lehetőséget ad arra, hogy bizonyos esetekben tőkekövetelmény alatt a várható értékkel csökkentett kockázatotott értéket értsük. A VaR pedig nem más, mint egy *kvantilis*, egy  $1 - \alpha$  rendű kvantilis: megmutatja azt az összeget, amelynél nagyobb teljes veszteség bekövetkezésének valószínűsége  $\alpha$ , azaz  $1 - \alpha$  biztonsággal mondhatjuk, hogy a vizsgált időszakban a teljes veszteség nem fogja meghaladni a VaR által megadott értéket. A fentiek alapján a következő módon adhatjuk meg a VaR precíz definícióját:

$$VaR_{1-\alpha}(S) = \sup\{x \in R \mid F_S(x) = P(S < x) < 1 - \alpha\}.$$

Ebben az esetben tehát az alsó kvantilis adja a VaR értékét, ezért  $VaR_{1-\alpha}$ -t az  $1 - \alpha$  rendű alsó VaR-nak szokás nevezni. Természetesen hasonlóan definiálható a felső VaR fogalma is a megfelelő felső kvantilis segítségével:

$$VaR^{1-\alpha}(S) = \inf\{x \in R \mid F_S(x) = P(S < x) > 1 - \alpha\}.$$

Az alsó és felső kvantilis nem szükségképpen esik egybe<sup>12</sup>. Abszolút folytonos eloszlásoknál azonos értéket ad adott szint mellett, hiszen ekkor egyszerűen legyen  $q$  az az érték, amelyre

$$F_S(q) = 1 - \alpha$$

teljesül, hiszen ekkor nyilvánvalóan

$$VaR_{1-\alpha}(S) = VaR^{1-\alpha}(S) = q.$$

Közismert, hogy a két érték különbözhet, például diszkrét eloszlások esetén bizonyos szinteken. (A precizitás kedvéért jegyezzük meg azt is, hogy bizonyos esetekben az sem teljesül, hogy pontosan  $\alpha$  és  $1-\alpha$  a  $VaR_{1-\alpha}$ -nál vagy a  $VaR^{1-\alpha}$ -nál nagyobb illetve kisebb veszteségek valószínűsége.) A VaR ezen tulajdonságaival ebben az írásban nem foglalkozunk részletesen, hiszen a későbbiekben tárgyalandó példák és problémák szempontjából nem lényeges. (Az érdeklődő olvasónak ajánljuk Gáll et. al. [2006] jegyzetét, vagy Szegő [2004] könyvét, ahol az említett kérdések részletesen vannak vizsgálva.) Megjegyezzük, hogy az alsó és felső VaR mellett más VaR fogalmakat is szokásos definiálni (például a fenti VaR fogalmak bizonyos átlagát véve), de ennek írásunk szempontjából nincs jelentősége.

A továbbiakban a fentiekben leírt módon meghatározott tőkekövetelményt operációs kockázatnak fogjuk nevezni az egyszerűség kedvéért. Működési kockázatok esetén a szóban forgó értékre a VaR elnevezés helyett szokásos a *Capital at Risk* (CaR) elnevezés is, amelyet talán *kockázatosított tőkének* nevezhetnénk.

A bázeli ajánlások szerint (Basel Committee on Banking Supervision [2001]) az összes kialakított osztály tőkekövetelményének összegeként adódik a pénzintézet operációs kockázatára vonatkozó teljes tőkekövetelménye. Tehát ha adott  $M$  osztály, melyek teljes vesztesége az adott időszakban rendre  $S^{(j)}$ , ahol  $j = 1, 2, \dots, M$ , és az ezekhez tartozó tőkekövetelmények értéke rendre

$$T^{(j)} = VaR_{1-\alpha}(S^{(j)}),$$

---

<sup>12</sup> Már egy egyszerű (akár kétértékű) diszkrét valószínűségi változónál is előfordulhat ilyen eset, hiszen az eloszlásfüggvénye lépcsős, ám az összetett veszteségek esetén ennek jelentősége nem igazán nagy egy megfelelően 'gazdag' modell esetén. Ezen problémakört példákkal együtt tárgyalja Gáll és Pap (2005).

akkor az LDA alapján a pénzügyi intézmény adott időszakhoz tartozó operációs kockázatokra meghatározott *teljes tőkekövetelménye*<sup>13</sup>

$$T = \sum_{j=1}^M T^{(j)} .$$

Jegyezzük meg azt is, hogy egyes esetekben a szabályozó lehetőséget ad arra is (bizonyos feltételek teljesülése mellett<sup>14</sup>), hogy a hitelintézet által képzett tőke csupán az ún. *nem várt veszteségekre* nyújtson fedezetet, ebben az esetben a következő összefüggés érvényes:

$$T^{(j)} = VaR_{1-\alpha}(S^{(j)}) - ES^{(j)} ,$$

ahol  $E$  a várható értéket jelöli.

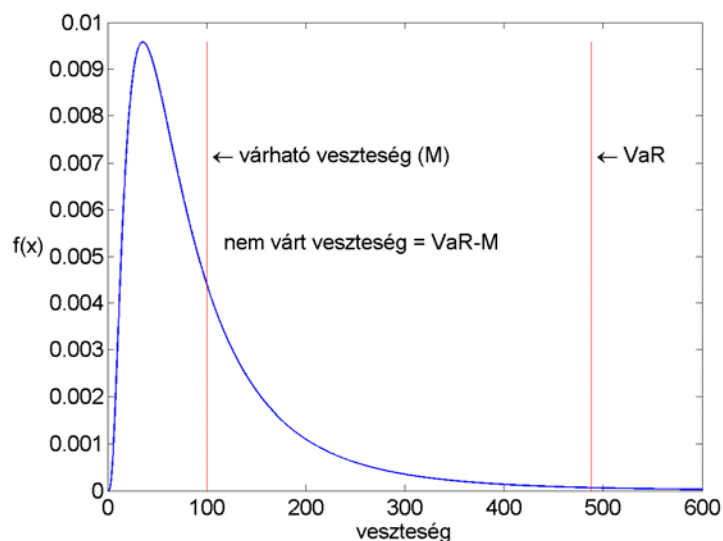
Az 1. ábra egy lognormális eloszláson<sup>15</sup> szemlélteti a VaR, az elvárt veszteség (várható veszteség) és a nem várt veszteség viszonyát a sűrűségfüggvénye mellett. A 2. ábra ugyanezen mennyiségek viszonyát mutatja egy összetett eloszlás, azaz a teljes veszteségek esetén.

---

<sup>13</sup> Frachot et. al. [2004] megmutatták, hogy valójában a tőkekövetelmények egy ilyen összegzése akkor indokolt elméletileg, ha a különböző kockázatokból származó veszteségek egyfajta teljes függősége ('tökéletes korrelációja') fennáll. Ez nem életszerű helyzetet tükröz, ugyanakkor az intézmények ettől csak akkor térhetnek el, ha a korreláció mérésére alkalmazott módszereit a felügyeleti szerv jóváhagyja. A veszteségek közötti függőségek kérdésével ezen tanulmány nem foglalkozik, az érdeklődő olvasónak ajánljuk Frachot et. al. [2004] írását.

<sup>14</sup> A várt (várható) veszteséget nem kell a tőkeképzés során figyelembe vennie az intézménynek, amennyiben bizonyítja a felügyelet felé, hogy szabályzatában meghatározta a várható veszteségek mérséklésére vonatkozó eljárásokat, azaz belső üzletviteli eljárásaiban azokat más módon (pl. céltartalék képzésben, termékei árazásakor) már figyelembe veszi.

<sup>15</sup> A dolgozat során gyakran dolgozunk lognormális eloszlással az egyes példák, problémák kapcsán. Ennek oka, hogy Basel Committee on Banking Supervision [2011] szerint ez a leggyakrabban alkalmazott eloszlás a működési kockázati tőke modellezésére.



3. ábra: Várható veszteség, kockázatos érték és a nem várt veszteség, mint statisztikai jellemzők illusztrálása a lognormális eloszláson ( $\mu=4,26$   $\sigma=0,83$ ).

Természetesen adódik az a kérdés, hogy milyen  $\alpha$  érték mellett érdemes a VaR-t számolni, és így a tőkekövetelményt meghatározni. Az uniós és magyar szabályozás egyaránt  $\alpha = 0,001$  mellett írja elő a VaR meghatározását egyéves időszak során bekövetkező működési kockázati veszteségekre vonatkozóan (200/2007. (VII. 30.) Korm. r.), azaz 99,9 százalékos biztonság mellett kell majd meghatározni a tőkekövetelményt operációs kockázatok esetén az AMA (így az LDA) módszertant választóknak. Ez természetesen nem jelenti azt, hogy ezt az egyetlen értéket tanácsos meghatározni. A VaR napjainkban a leggyakrabban használatos kockázati mérték a modern pénzügyben<sup>16</sup>. Annak meghatározása különböző  $\alpha$  értékek esetén fontos információt ad a vállalt kockázatokról, ezért érdemes néhány VaR értéket meghatározni különböző biztonsági szinteken még akkor is, ha azokat a készülő szabályozás nem követeli meg, hiszen a veszteségek eloszlásának tulajdonságai, különösképpen a farok (szél) viselkedése (nagy veszteségek) jól tükröződik a különböző VaR értékekben.

Fontos megjegyeznünk azt is, hogy portfóliók, pénzügyi eszközök esetén nem a VaR az egyetlen kockázati mérték (mutató), amelyet lehet, vagy ajánlatos használni<sup>17</sup>. Sőt, a

<sup>16</sup> A Riskmetrics módszertan és a Basel I-es kiegészítés piaci kockázatok tekintetében nagymértékben segítette a VaR, mint kockázati mérték elterjedését.

<sup>17</sup> A 2008-as válság következtében felgyorsult szabályozói folyamatok - Basel 2.5 és Basel III javasolja az Expected Shortfall kockázati mérték bevezetését is, elsődlegesen piaci kockázatok mérésére. Az új szabályozói irányelvek foglalkoznak a modellek fogyatékoságaival is, erre példa, hogy I. pillérben a korábbi VaR alapú kereskedési könyvi tőkekövetelményt Stress VaR és Incremental Risk Charge egészíti ki.

szakirodalomban számos publikációt találhatunk arról, hogy az egyes mutatóknak milyen előnyös és hátrányos tulajdonságaik vannak. A Value at Risk esetén a legfontosabb kritika az, hogy nem teljesíti a szubadditivitás tulajdonságát<sup>18</sup> és nem ragadja meg kellően az eloszlás szélén található nagy veszteségeket. Előbbi miatt a VaR nem egy koherens kockázati mérték. Jelen írásban nem célunk a kockázati mértékek és a VaR tulajdonságainak részletes tárgyalása, az érdeklődő olvasó többet tudhat meg ezen kérdéskörrel többek között Acerbi [2004], Delbaen [2000] és Gáll et. al. [2005] munkáiban, de itt említhetnénk számos más szakirodalmi forrást is.

## 5.2. Tőkekövetelmény számítás egyéb módszerei

### 5.2.1. Közelítéssel módszer

Fontos megemlíteni, hogy a működési kockázati tőke meghatározása során ugyan az AMA-t alkalmazó hitelintézetek közül 70% alkalmaz LDA alapú megközelítést, mellette más módszertanok is léteznek. Ezek közül az egyik legegyszerűbb az aggregált eloszlás (S) közelítése. Ennél a módszernél az aggregált káreloszlást közvetlenül közelítjük egy előre választott eloszlással. A módszer használata esetén a gyakoriság és egyedi veszteség eloszlások néhány jellemzőjét kell csupán megbecsülnünk, melyeket az összetett eloszlás várható értékének és varianciájának a meghatározásához használunk fel. Az aggregált eloszlás közelítésére alkalmazott tipikus eloszlások a normális és a lognormális eloszlás.

Jelölje továbbra is S összetett eloszlást:  $S = X_1 + X_2 + \dots + X_\eta$ , ahol  $X_i$  jelöli az egyedi veszteségeket,  $\eta$  pedig az események bekövetkezési gyakoriságát. Az S aggregált veszteség eloszlás centrális momentumai felírhatók az egyedi veszteségeloszlás és gyakoriságeloszlás momentumainak segítségével:

$$ES = EX \cdot E\eta, \quad (2)$$

$$D^2S = E\eta \cdot D^2X + D^2\eta \cdot (EX)^2, \quad (3)$$

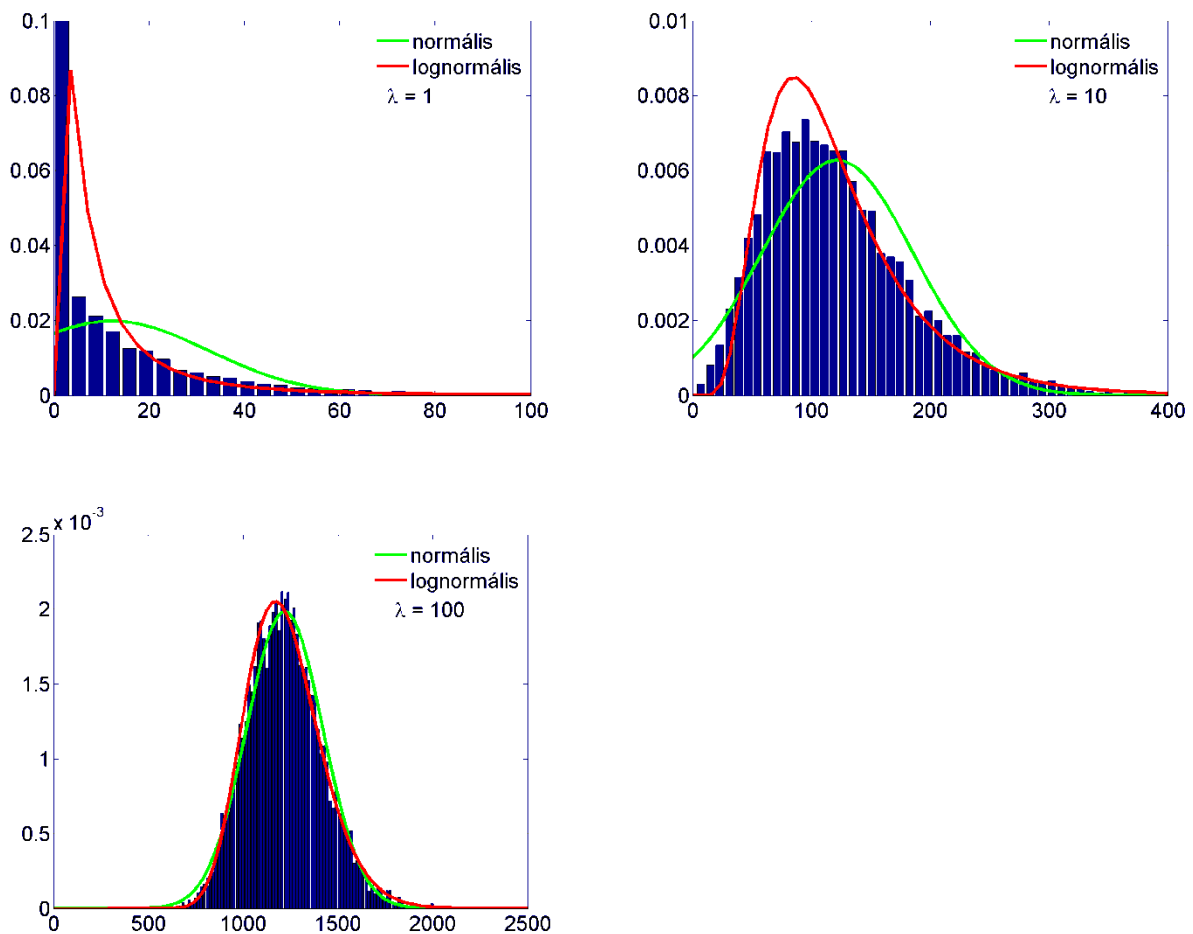
ahol  $E$  és  $D^2$  rendre várható értéket és varianciát (szórásnégyzetet) jelöl.

---

<sup>18</sup> Szubadditívnek nevezünk egy kockázati mértéket, ha két portfólió esetén a portfóliókra külön-külön számolt kockázati mértékek -azaz kockázatok- összegénél nem lehet nagyobb a két portfólió egyesítésével létrehozott portfólió kockázati mértéke, azaz kockázata.

Miután a gyakoriság és a veszteségeloszlás szükséges jellemzőit megbecsültük, azokat a fenti egyenletek jobb oldalába behelyettesítettük, miközben az egyenletek bal oldalán az illeszteni kívánt kétparaméteres összetett eloszlás (esetünkben normális vagy lognormális) elméleti jellemzői (várható érték és variancia) állnak, az egyenletrendszer megoldásával már számíthatjuk is a választott eloszlás paramétereit (Klugman et. al. [2004]).

A centrális határeloszlás tétele alapján a gyakoriság eloszlás nagy várható értékére az összkár eloszlásának jó közelítése a normális eloszlás. Panjer [2006] alapján ez elsődlegesen a Poisson, binomiális vagy negatív binomiális gyakoriság eloszlások esetén teljesül. A kárgyakoriság kicsi várható értéke esetén viszont az összetett káreloszlásra az aszimmetrikus eloszlások adhatnak jó közelítést, ilyenkor alkalmazhatjuk pl. a lognormális eloszlást. Természetesen az így illesztett eloszlás VaR értéke fogja jelenti az adott kockázati osztályra jellemző tőkét. Részletesebb eredményekért lásd (Nagy et. al. [2009]) publikációját.



4. ábra:  $\lambda = [1, 10, 100]$  paraméterű Poisson gyakoriságokból és a  $\alpha=4,8$  és  $\beta=46$  paraméterű Pareto egyedi veszteség eloszlásból generált aggregált eloszlás közelítése normális és lognormális eloszlásokkal.



### 5.2.2. Single Loss Approximation

Az egyedi veszteséeloszlás modellezése szempontjából gyakran vastag szélű (heavy tail), vagy semi heavy tail eloszlások jönnek szóba. Ezek egy része (pl. lognormális, Weibull, Pareto, log-gamma eloszlások) a subexponenciális eloszláscsaládba tartozik. Ezen eloszláscsalád tagjainak fontos tulajdonsága a következő:

$$\lim_{x \rightarrow \infty} \frac{P(X_1 + \dots + X_n > x)}{P(\max(X_1, \dots, X_n) > x)} = 1, \quad n \geq 2 \text{ esetén.} \quad (4)$$

Ez azt jelenti, hogy az  $n$  darab veszteség összege valószínűleg nagy lesz, hiszen jó eséllyel lesz közöttük egy olyan hatalmas veszteség, ami dominálja a szummát. Tehát a veszteségek összege nem a sok kicsi felhalmozott veszteség miatt ér el egy kritikus küszöböt, hanem egyetlen nagy veszteség (single loss) miatt.

Böcker et. al. [2005] megmutatta, hogy a működési kockázati tőkekövetelmény, avagy az aggregált eloszlás megfelelő kvantilise subexponenciális eloszlások esetén analitikusan is meghatározható. Szubexponenciális egyedi veszteséeloszlás ( $F$ ) esetén

$$VaR(p) = F^{\leftarrow} \left( 1 - \frac{1-p}{EN} (1 + o(1)) \right), \quad p \rightarrow 1. \quad (5)$$

Ebből adódóan a Single Loss Approximation szerinti kvantilis érték:

$$VaR(p) = F_s^{-1}(p) \approx F^{-1} \left( 1 - \frac{1-p}{E(N)} \right). \quad (6)$$

A fenti képlet direkt kapcsolatot teremt az aggregált eloszlás ( $F_s$ )  $p$  konfidenciaszinthez tartozó kvantilise és az  $F$  egyedi veszteséeloszlás  $1 - \frac{1-p}{E(N)}$  konfidenciaszinthez tartozó kvantilise között. A tőke meghatározása csupán az egyedi veszteséeloszlásból történik a gyakoriság várható értékének felhasználásával. A Single Loss Approximation módszertanára később még visszatérünk, itt csupán a dolgozat során felmerülő problémák megértéséhez nyújtottunk előzetes áttekintést.

## 6. Az operációs kockázat becslése, és kapcsolódó problémák

Az előzőekben megismerkedtünk azzal, hogy LDA esetén miként érdemes az operációs kockázatot jellemezni és így magát a tőkekövetelményt megadni. Amint láttuk, annak megadása feltételezi a veszteségeloszlás és a gyakoriságeloszlás ismeretét. Azok együttesen már meghatározzák az teljes veszteség eloszlását, amely pedig nyilvánvalóan meghatározza a kérdéses VaR értékeket is. A gyakorlatban természetesen nem ismertek az említett (elméleti) eloszlások, így a VaR értékeket becsülni szükséges valamilyen statisztikai módszerrel. A VaR becslése egyszerű feladatnak tűnik, hiszen egy kvantilis becsléséről van szó, ám számos statisztikai kérdés felvetődik a becslés módszerének megválasztása során és a becslés során. A dolgozat ezen részében ezeket a kérdéseket kívánjuk áttekinteni.

Mivel egy kvantilis becslése a feladatunk, így természetesen adódik egy közvetlen módszer, nevezetesen tekintsük a statisztikában jól ismert becslését a kvantilisnek, úgy is mondhatnánk, hogy az empirikus kvantilist. (Itt most nem térünk ki arra, hogy mi a precíz definíciója a szakirodalomban ismert becsléseknek, s melyek a becslések statisztikai tulajdonságai, különös tekintettel a különböző kvantilis fogalmak okozta apróbb különbségekre.) Ezt nevezhetjük egy nemparaméteres módszernek is, hiszen valójában az nem feltételezné az eloszlások és azok paramétereinek ismeretét, becslését. Ehhez mindössze a teljes veszteségeket tartalmazó mintára lenne szükség minél több megfigyelt időszakra, azaz minél nagyobb mintaelemszámmal.

Azonban ez nyilván nem járható út működési kockázatok esetén, hiszen a teljes veszteségadatok száma, azaz a megfigyelt időszakok száma nagyon kevés a magyar pénzügyintézeteknél, tipikusan néhány év. Ehhez jegyezzük meg, hogy az operációs kockázatok számításához szükséges adatbázisok következetes kialakítása a legtöbb pénzügyintézetnél néhány, esetenként körülbelül mindössze 10 évre nyúlik vissza. De ha el is játszunk egy pillanatra a gondolattal, hogy mi lenne, ha egy sok megfigyelt évet (időszakot) tartalmazó mintánk lenne, vagy esetleg több évtizednyi minta állna rendelkezésünkre, akkor is láthatnánk, hogy ez a közvetlen kvantilisbecslés statisztikailag nem igazán adna megbízható eredményt. Itt két fontos dologra hívjuk fel az olvasó figyelmét. Az egyik probléma, hogy több év adatainak használata során egyáltalán nem lehetünk biztosak abban, hogy a teljes

veszteséget leíró eloszlások nem módosultak, így a minta azonos eloszlásúsága csorbul, ami számos problémát vetne fel. Másrészt jegyezzük meg, hogy tipikusan 99,9%-os biztonsági szinthez, tehát igen magas szinthez akarunk VaR-t becsülni, amelynél nem engedhetjük meg azt, hogy csak a teljes veszteségeloszlásokat tartalmazó mintát használjuk, elveszítve ezzel rengeteg információt az egyedi veszteségek és a gyakoriság eloszlására vonatkozóan. Másképpen úgy is megfogalmazhatjuk ezt a problémát, hogy az intézmény rendelkezésére álló minta (teljes veszteségadatok száma) időszakonként (évente) csupán egy elemmel bővül, így több évtizednyi adatgyűjtés után is egy néhány tucatnyi elemet tartalmazó mintából kellene meghatározni egy nagyon magas konfidenciaszinthez tartozó empirikus kvantilist.

A tőkekövetelmény becsléséhez így a szakirodalom (és a tipikus banki gyakorlat) elsődlegesen egy paraméteres, közvetett utat (LDA) javasol. Ennek lényege az, hogy a teljes veszteségeket felépítő egyedi veszteségek eloszlását és a gyakoriság eloszlását próbáljuk meg meghatározni. Ez esetben adott eloszláscsaládok paramétereinek becslését kell elvégeznünk, majd abból következtetni a teljes veszteség eloszlására és annak kvantiliseire. Megjegyezzük, hogy az egyedi veszteségeloszlás mintaelemeit felhasználva is limitáltak a lehetőségeink a működési kockázati tőke helyes meghatározása során, melyet a már felvezetett Single Loss Approximation keretein belül fogunk megmutatni.

Itt most összefoglaljuk röviden, hogy milyen feladatokat kell végrehajtanunk, ha körültekintően akarjuk elvégezni a tőkekövetelmény kiszámítását a felvázolt paraméteres megközelítés esetén. A szükséges lépések az alábbiak.

- Homogén minta létrehozása a megfelelő szűrési feltételek meghatározásával (időszak, veszteség kategóriák és üzletágak alapján az osztályok rögzítése).
- Az egyedi veszteségek modellezésére alkalmas eloszláscsaládok kiválasztása.
- A gyakorisághoz használt eloszláscsaládok rögzítése.
- A legjobban illeszkedő eloszláscsalád kiválasztása a vizsgált (veszteség)osztály esetén megfelelő statisztikai módszerekkel (modellszelekció).
- A veszteségek és a gyakoriság eloszlásai esetén a szükséges paraméterek becslése és a kialakított modell illeszkedésének vizsgálata.

- A teljes veszteség eloszlásának, jellemzőinek, mutatóinak (pl. momentumok) meghatározása, illetve becslése, különös tekintettel a szükséges kvantilis (azaz VaR) meghatározására vagy becslésére.
- Külső adatbázisok használatának megfontolása és a becslési módszerek korrekciója ilyen esetekre.

A felvázolt feladatok nem jelentenek sorrendet, sőt, egyes problémákat –mint például a paraméterek becslése és modellszelekció- gyakorta egyszerre tudjuk kezelni. A fenti felsorolás sokkal inkább azt hivatott kiemelni, hogy melyek azok a fontos lépések és szempontok, amelyeket a munkánk során mérlegelnünk kell, hogy egy megbízható, az ellenőrző szervek számára is elfogadható belső módszertant dolgozzunk ki egy adott pénzügyintézet működési kockázatainak modellezésére és a szükséges tőkekövetelmény meghatározására.

A következő részben a felsorolt lépések végrehajtásához szükséges (javasolt) módszereket, eljárásokat és kapcsolódó problémákat tekintjük röviden át. Nem célunk az egyes módszerek részletes elméleti bemutatása, hiszen azok közismertek a statisztikai irodalmában. Célunk viszont egyrészt a szakirodalomban javasolt és alkalmazott módszerek áttekintése és azok alkalmazási kérdéseinek tisztázása, másrészt pedig egyes kérdések kiemelése, amelyek esetén úgy gondoljuk, hogy a tárgyalt operációs kockázati problémák specifikumaiból adódóan fontosabbak lehetnek, mint más alkalmazásoknál.

### *6.1. A mintáról*

Először röviden kitérünk a mintanagyság problémájára. Ezt a kérdést részben már az előzőek során érintettük, hangsúlyozva, hogy tipikusan néhány év adatai állnak egy pénzügyintézet rendelkezésére (és esetenként azok sem teljes körűen). Egy adott veszteségkategóriában pedig évente néhány tucat megfigyelés már átlagon felüli egy magyar pénzügyintézetnél (a pénzügyintézet mérete miatt), így egy adott veszteségkategórián belül 1-1 üzletágban pedig már csak néhány adat van évente, sőt, egyes esetekben akár előfordulhat, hogy egyetlen adat sincs.

Ennélfogva különösen megfontolandó, hogy milyen egységekre számolunk külön tőkekövetelményt. Akár az operációs kockázat mátrixának mind az 56 elemére -azaz üzletág-veszteségkategória kombinációjára- számolhatnánk tőkekövetelményt, majd azok összegeként a teljes tőkekövetelmény adódna. Azonban a minták rendkívül alacsony mérete miatt statisztikai szempontból ajánlott összevonásokat alkalmazni és így az 56-nál lényegesen kevesebb kockázati osztályt létrehozni. Például összevonhatunk üzletágak vagy veszteségkategóriák mentén. Azt is meg kell azonban gondolnunk, hogy az összevonások után még reálisnak tartjuk-e a teljes eloszlásra felírt összetett modellt, amelyben például feltételeztük az egyedi veszteségek eloszlásának azonosságát egy adott osztályon belül. Így talán ajánlatosabb a veszteségkategóriákat nem összevonni, vagy csak bizonyos kategóriákat összevonni, míg talán üzletáganként indokoltabb lehet az összevonás. A gyakorlatban is tipikusan használt, kézenfekvő megoldás az üzletágak összevonása, azaz csak a veszteségkategóriák szerinti bontás alkalmazása.

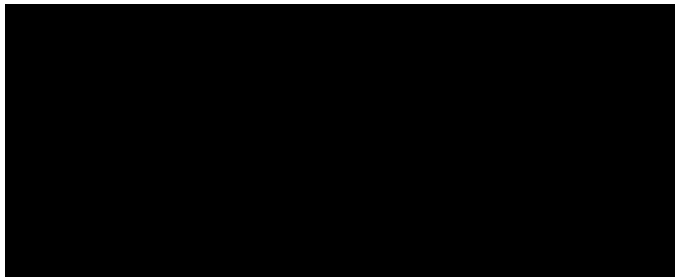
Fontos megemlítenünk, hogy itt nem adhatunk egy általános ajánlást. Csak az adott pénzügyi intézet saját mintáját, korábbi évek tapasztalatát és az adatokon végzett statisztikai vizsgálatokat figyelembe véve érdemes és lehet kialakítani egy belső rendszert az összevonásokra. Hiszen például a termékinálat, az ügyfélkör sajátosságai lényegesen befolyásolhatják az egyes kategóriákba eső adatok számát és jellegét (eloszlását). Továbbá fontos azt is hangsúlyozni, hogy a szabályzónak is elfogadható kell, hogy legyen egy kialakított belső modell.

### *6.1.1. Az optimális mintaelemszámról*

Számos szerző foglalkozik azzal a kérdéssel, hogy mekkora a VaR (tőke) becslés hibája, illetve mi lenne az optimális mintaelemszám, amely mellett a tőke már elfogadható hibával becsülhető. Szabályozói szemszögből a VaR alulbecslése, míg tulajdonosi szemmel nézve a tőke túlbecslése jelenti a problémát. A becslés hibájával később foglalkozunk, arra számos pontban visszatérünk, itt csak ízelítőt adunk az optimális mintaelemszám nagyságrendjéről. Az egyedi veszteségek éves gyakorisága és eloszlásának tükrében Frachot et. al. [2004] szimulációs kísérletek segítségével ad választ a tőke becslési hibájára és optimális mintaelemszámra. Elemzésében, arra a kérdésre ad választ, hogy mekkora

mintaelemszám szükséges ahhoz, hogy a becült tőke ( $VaR_{EST}$ ) a tényleges tőke ( $VaR_{TRUE}$ )  $1-c$  százalékánál  $\alpha$  valószínűséggel nagyobb legyen?

Eredményét az optimális, vagy inkább az elfogadható mintanagyság tekintetében  $\lambda=100$  paraméterű Poisson gyakoriság eloszlás és lognormális veszteség eloszlás különböző  $\mu, \sigma$  paramétereit mellett az alábbi táblázatok foglalják össze.



1. táblázat: optimális mintaelemszám  $\lambda=100, \mu=9, \sigma=2$  esetén. Az adatgyűjtési küszöb 0 euró.

Egy átlagosan évi 100 adatot gyűjtő bank esetében 6 260 elemű mintára lenne szükség ahhoz, hogy 90%-os valószínűséggel ne becsülje több mint 10 százalékkal alá a működési kockázati tőkét.

c	$\alpha$			
	60%	70%	80%	90%
10%	400	1740	4500	10430
20%	100	430	1120	2600
30%	40	190	500	1150
40%	20	100	280	650
50%	10	60	180	410

2. táblázat: optimális mintaelemszám  $\lambda=100, \mu=9, \sigma=2$  esetén. Az adatgyűjtési küszöb 5.000 euró. Egy átlagosan évi 100 adatot gyűjtő bank esetében több mint 10 000 elemű mintára lenne szükség ahhoz, hogy 90%-os valószínűséggel ne becsüljük több mint 10 százalékkal alá a működési kockázati tőkét.

c	$\alpha$			
	60%	70%	80%	90%
10%	410	1760	4540	10530
20%	100	440	1130	2630
30%	40	190	500	1170
40%	20	110	280	650
50%	10	70	180	420

3. táblázat: optimális mintaelemszám  $\lambda=100$ ,  $\mu=7$ ,  $\sigma=2.5$  esetén. Az adatgyűjtési küszöb 0 euró.

Egy átlagosan évi 100 adatot gyűjtő bank esetében 10 530 elemű mintára lenne szükség ahhoz, hogy 90%-os valószínűséggel ne becsülje több mint 10 százalékkal alá a működési kockázati tőkét.

c	$\alpha$			
	60%	70%	80%	90%
10%	820	3530	9110	21120
20%	200	880	2270	5280
30%	90	390	1010	2340
40%	50	220	560	1320
50%	30	140	360	840

4. táblázat: optimális mintaelemszám  $\lambda=100$ ,  $\mu=7$ ,  $\sigma=2,5$  esetén. Az adatgyűjtési küszöb 5.000 euró.

Egy átlagosan évi 100 adatot gyűjtő bank esetében 10 530 elemű mintára lenne szükség ahhoz, hogy 90%-os valószínűséggel ne becsülje több mint 10 százalékkal alá a működési kockázati tőkét.

A fenti táblázatokból látható, hogy tízezres nagyságrendű mintára van szükség ahhoz (miközben a banknak várhatóan csak 500 van 5 év után is a fenti paraméterek mellett!), hogy a hitelintézet becsült tőkéje 90 százalékos valószínűséggel ( $\alpha$ ) ne legyen kevesebb az igaz tőkeérték 90 (1-c) százalékánál.

Pareto eloszlások ( $F(x) = 1 - x^{-\alpha}$ ,  $\alpha > 0, x \geq 1$ ) esetén a szükséges mintaelemszám meghatározásához a kvantilis sztenderd hibájára -  $\sigma(q_p)$  vonatkozó becslést (Kiefer [1967]) használja fel (Cope et. al. [2009]), miszerint

$$\sigma(q_p) \approx \frac{1}{f(q_p)} \sqrt{\frac{p(1-p)}{n}}. \quad (7)$$

10 százalékos relatív hiba<sup>19</sup> eléréséhez  $p=0,999$ -es valószínűséghez tartozó kvantilis becslése során  $\alpha=1,2$  paraméter mellett, 277 000 elemű mintanagyságot határoztak meg.

<sup>19</sup> A relatív hibát a  $\frac{2\sigma(q_p)}{q_p}$  formában definiálták.

Láthatjuk, hogy a különböző szerzők más és más mintanagyságot javasolnak, vagy határoznak meg (természetesen nem teljesen ugyanabban a kontextusban – más gyakoriság és egyedi veszteség eloszlás feltételezése mellett), ugyanakkor a közös vonás bennük, hogy a tőke relatíve pontos meghatározásához nagyszámú mintára van szükség, miközben a hitelintézetnek ez nem áll rendelkezésére. A szükséges minta nagyságát részben a veszteséggeneráló folyamat határozza meg, azaz milyen gyakorisággal és mennyire vastag szélű eloszlásból származnak az adatok, ugyanakkor nem szabad megfeledkeznünk a szabályozó által determinált tőkekövetelmény definícióról sem (éves időtávon, 99,9%-os valószínűséggel kell a tőkének fedezetet nyújtania a veszteségekre)! Ez utóbbi fontos következménye, hogy a hitelintézet sosem rendelkezhet a megfelelő mintaelemszámmal tőkéjének a becsléséhez!

Tegyük fel, hogy évente átlagosan  $k$  darab veszteséget figyelünk meg. Ahhoz, hogy a szabályozó által meghatározott tőkekövetelményt közvetlenül meg tudjuk határozni 1 000 év aggregált veszteségét kell megfigyelnünk, ekkor tudjuk a tőkét 99,9%-os biztonsággal közvetlenül meghatározni. Ez számunkra azt jelenti, hogy az egyedi veszteség eloszlásból  $k \times 1000$  darab megfigyeléssel kell rendelkezünk, tehát az intézménynek 1 000 év adatára van szüksége. Ehhez képest az intézmény mindössze 5 évnyi  $k \times 5$  darab veszteséggel rendelkezik a tőkeszámításhoz.

A működési kockázati veszteségek eloszlása a szubexponenciális eloszláscsaládba tartozik, ami azt jelenti, hogy az éves aggregált veszteséget tipikusan 1 veszteség, a maximális veszteség határozza meg. Ebből adódik, hogy az egyedi veszteségeloszlás viselkedése a szélső tartományban fogja meghatározni az aggregált eloszlás VaR-ját, a tőkét.

A 99,9%-os konfidenciaszint melletti VaR meghatározása 5 éves idősről a veszteségek  $1000/5=200$ -szoros extrapolálását, és egyben parametrikus modell használatát teszi szükségessé. Az eloszlás extrapolálása olyan tartományba (a becsülni kívánt kvantilis tartományába), ami nem megfigyelhető és ezáltal visszamérhető az intézmény számára, óriási bizonytalanságot rejt magában. Ezt a problémát a későbbiekben extrapolációs problémaként hivatkozzuk (Cope et. al. [2009]).

### *6.1.2. Single Loss Approximation és mintanagyság*

Amint korábban megmutattuk az SLA keretében a tőke az alábbi módon határozható meg:



$$VaR = F_s^{-1}(p) \approx F^{-1}\left(1 - \frac{1-p}{E(n)}\right). \quad (8)$$

Ez direkt kapcsolatot teremt az aggregált eloszlás ( $F_s$ )  $p$  konfidenciaszinthez tartozó kvantilise és az egyedi veszteségeloszlás ( $F$ )  $1 - \frac{1-p}{E(n)}$  konfidenciaszinthez tartozó kvantilise között.

A következő táblázatban  $E(n)$  néhány értékére / nagyságrendjére összefoglaljuk, hogy milyen egyedi veszteségeloszlás konfidenciaszintet is jelentenek ezek számunkra.

Éves gyakoriság várható értéke – $E(n)$	Egyedi veszteség eloszlás konfidenciaszintje - $1 - \frac{1-p}{E(n)}$	Direkt kvantilis becsléshez szükséges adatok minimális száma	Működési kockázati veszteségek várható száma 5 év alatt
0,1	0,99	100	0,5
1	0,999	1000	5
10	0,9999	10 000	50
100	0,99999	100 000	500
1.000	0,999999	1 000 000	5 000

5. táblázat: Események éves száma, és az egyedi veszteségeloszlás konfidenciaszintje néhány kiemelt értékre. Látható, hogy minél nagyobb az események éves száma, annál távolabbi kvantiliset kell becsülni az egyedi veszteség eloszlásnak a tőke meghatározásához. Ezentúl a táblázat azt is tartalmazza, hogy mi lenne az adatok minimális száma egy direkt kvantilis becsléshez.

A táblázat minden rekordja esetében jól látható, hogy lényegesen (egészen pontosan 200-szor) több egyedi veszteség adatra lenne szükség az extrapolációs probléma elkerüléséhez, mint amennyivel egy pénzüintézet 5 év alatt rendelkezik.

A szabályozó ugyan definiál két olyan modellkomponenst (külső adatok és scenárió elemzés), amit a belső modellnek tartalmaznia kell, és segíthet a becslés javításában, ugyanakkor azok használata esetén is számos probléma merül fel. A külső adatok kapcsán

abban reménykedhetnénk, hogy az adatkonzorcium még másik 199 hasonló intézmény adatait tartalmazza, és ezáltal megkerülhetjük az extrapolációs problémát. Ez sajnos még a nagy nemzetközi adatkonzorciumok (pl. ORX) esetében sem teljesül, illetve felmerül a kérdés, hogy az adatkonzorcium tagjainak veszteséggeneráló folyamata mennyiben hasonlít az adott intézményéhez. Szenárióelemzés keretében a jó szakértők adhatnak reményt, akik a kis gyakoriságú extrém veszteségeket tudják jól megbecsülni (Hajnal et. al. [2007]).

## 6.2. A gyakoriság eloszlásról

### 6.2.1. Tipikus gyakoriság eloszlások

A gyakoriság eloszlások esetén mondhatjuk, hogy tekinthetünk tetszőleges nemnegatív egészértékű eloszlást. Ezek közül is elsősorban a 6. táblázatban szereplő három eloszlás (Poisson, binomiális és negatív binomiális) a leginkább javasolt. Ezeket akár tekinthetjük egy eloszláscsaládnak is, hiszen mindegyik ún.  $(a, b, 0)$  típusú eloszlás, amely azt jelenti, hogy léteznek olyan valós  $a, b$  paraméterek, hogy az  $\eta$  gyakoriság eloszlására teljesül az alábbi rekurzió:

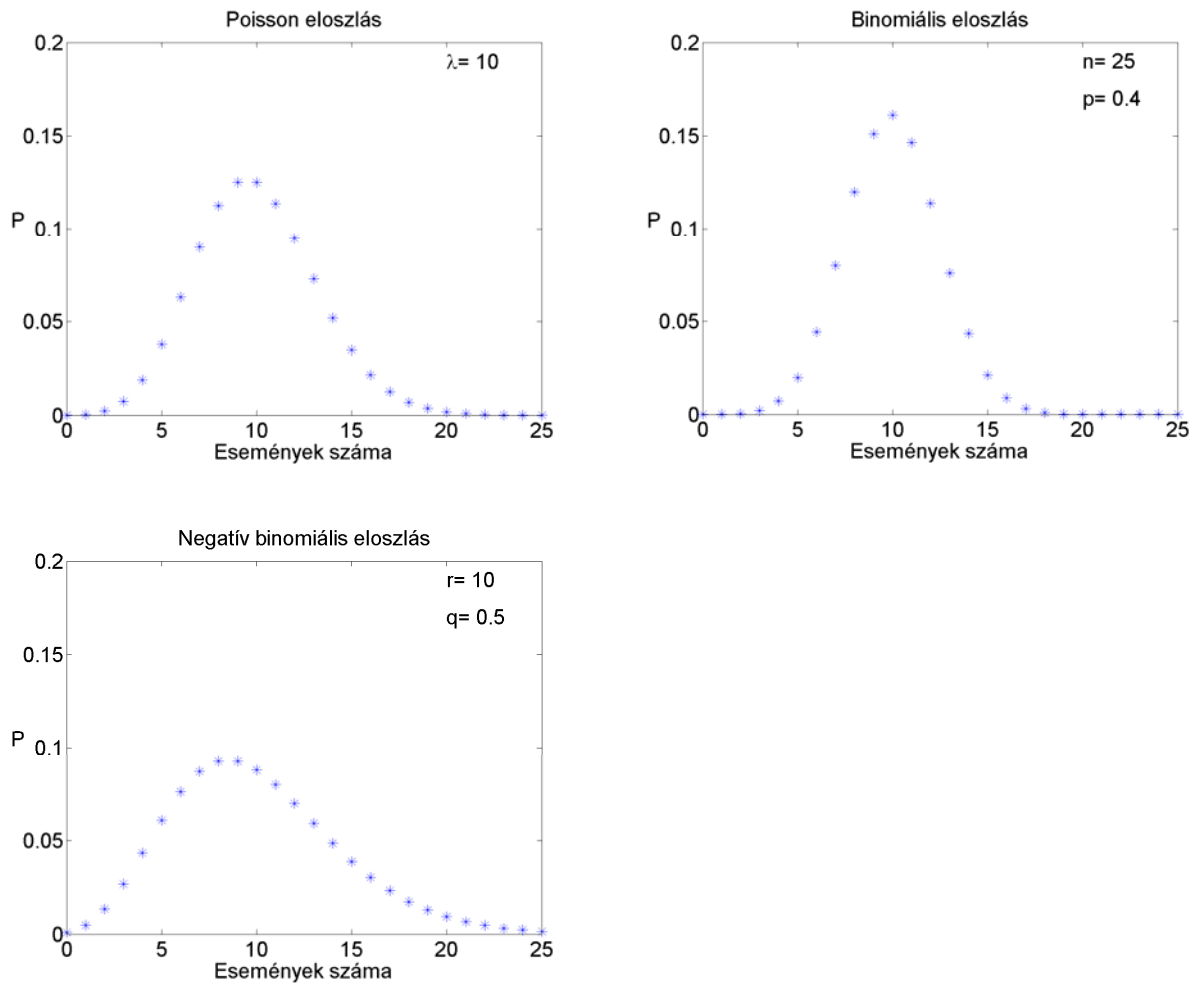
$$P(\eta = n) = \left(a + \frac{b}{n}\right)P(\eta = n - 1), \quad (9)$$

minden pozitív egész  $n$  esetén. Könnyen bizonyítható, hogy egy (nemnegatív egészértékű) diszkrét eloszlás pontosan akkor  $(a, b, 0)$  típusú, ha az Poisson, binomiális vagy negatív binomiális.

Eloszlás	Paraméterek	$P(\eta=k)$	$E(\eta)$	$D^2(\eta)$	a,b értékek
Binomiális	n: pozitív egész, $0 \leq p \leq 1$	$\binom{n}{k} p^k (1-p)^{n-k}$ , $k = 0, 1, \dots, n.$	$np$	$np(1-p)$	$a = -\frac{p}{1-p}$ $b = \frac{(n+1)p}{1-p}$
Poisson	$\lambda \geq 0$	$\frac{\lambda^k e^{-\lambda}}{k!}$ $k = 0, 1, \dots$	$\lambda$	$\lambda$	$a = 0$ $b = \lambda$

Negatív binomiális	$r$ : pozitív, $0 \leq q < 1$	$\frac{\Gamma(r+k)}{\Gamma(r)k!} (1-q)^r q^k$ , $k = 0, 1, \dots$	$\frac{rq}{1-q}$	$\frac{rq}{(1-q)^2}$	$a = q$ $b = (r-1)q$
--------------------	----------------------------------	--	------------------	----------------------	-------------------------

6. táblázat: (a,b,0) típusú eloszlások.



5. ábra: A Poisson ( $\lambda=10$ ), binomiális ( $n=25$ ,  $p=0,4$ ) és negatív binomiális ( $r=10$ ,  $q=0,5$ ) eloszlások szemléltetése. Az eloszlások várható értéke egyaránt 10. Az ábrákon látható, hogy habár a várható értékek megegyeznek, az eloszlások szóródása különböző.

### 6.2.2. Gyakoriság eloszlás választás

Működési kockázatok esetén különösen kicsi a gyakoriságokhoz tartozó minták elemszáma, hiszen egy-egy megfigyelt időszak csak egy újabb mintaelemet ad. Ennélfogva -

bár matematikailag könnyen kezelhető lenne mindhárom említett eloszlás, sőt az azok közötti választás is egyértelmű<sup>20</sup>, és ahogy később látni fogjuk, ezen eloszlások esetén jól használható rekurziós képlet is ismert az együttes eloszlás meghatározására- azt javasolhatjuk általánosságban, hogy a mindössze egy paraméterrel rendelkező Poisson eloszlást érdemes csak használni operációs kockázatok esetén. (A dolgozatban ezért mi is csak Poisson eloszlású gyakoriságokkal végzünk szimulációs vizsgálatokat.) A kis mintaméret miatt az is szükséges lehet, hogy a pénzügyi belső szakértői (szubjektív) becslést használjon statisztikai módszerek (pontbecslés) helyett / mellett, ha a statisztikai becslések hibája elfogadhatatlanul nagyra bizonyul.

A Basel Committee on Banking Supervision [2011] felmérése alapján a hitelintézetek tipikusan a Poisson eloszlást választották a működési kockázatok modellezésére. Ennek okai a következők lehetnek:

- könnyű kezelhetőség;
- elméleti támogatottság;
- kedvező matematikai tulajdonságok:
  - az adatok korlátozott mennyiségben állnak rendelkezésre, emiatt egyszerű modellt érdemes választani. Poisson eloszlás esetén csak 1 paramétert kell becsülni,  $\lambda = E(N)$ .
  - zárt az összegképzésre  $P(a)+P(b)=P(a+b)$ , azaz független Poisson eloszlású változók összege is Poisson eloszlású marad, ráadásul az új eloszlás paramétere a komponensek paraméterének az összege. Ebből adódóan kedvező tulajdonságokkal rendelkezik az időtartományban, azaz pl. a negyedévenkénti Poisson eloszlású gyakoriságok könnyen konvertálhatók (aggregálhatók) éves gyakorisággá, elkerülve az esetleges éven belüli szezonális jelenségeket (ami pl. a negatív binomiális eloszlás varianciáját már növelné).

---

<sup>20</sup> Klugman [2004], valamint Arató [1997] alapján az alábbi hüvelykujj szabályt fogalmazhatjuk meg a gyakoriság eloszlások közötti választásra:

ha  $E\eta < D^2\eta$ , akkor negatív binomiális,

ha  $E\eta = D^2\eta$ , akkor Poisson,

ha  $E\eta > D^2\eta$ , akkor binomiális eloszlást választunk,

ahol  $E$  és  $D^2$  rendre várható értéket és varianciát jelölnek.

- o zárt a csonkolásra (küszöb alatti veszteségek nem jelentése az adatbázisba), lásd consistency theorem (Embrechts et. al. [1997]).

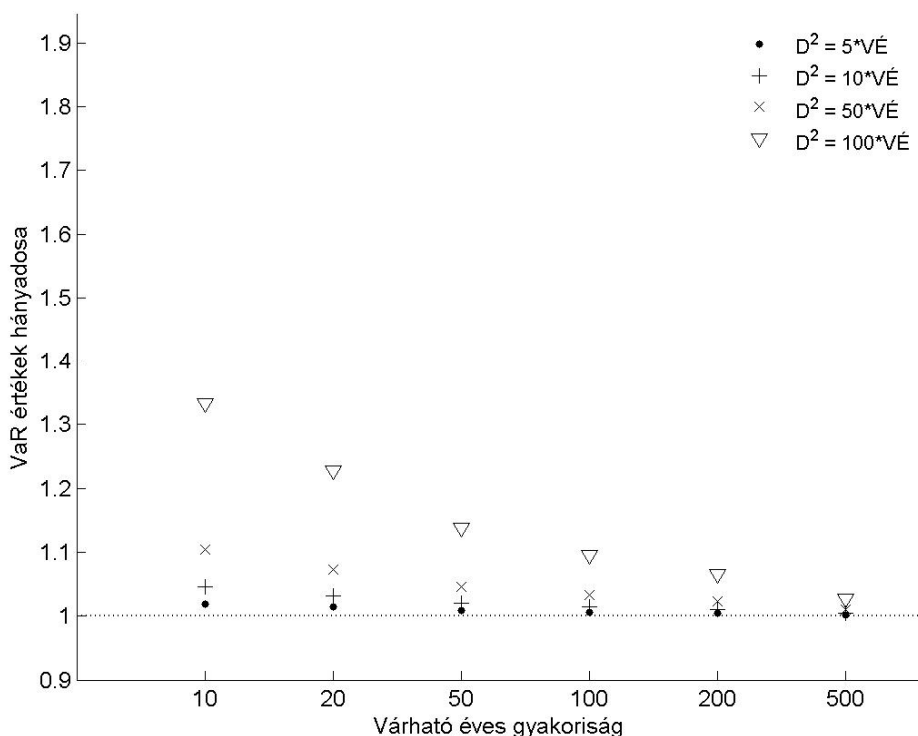
$$N_{X|X \geq a} \approx P(\lambda) \rightarrow N_X \approx P\left(\frac{\lambda}{1-F(a)}\right), \text{ ahol } F(a) = P(X \leq a), \text{ ahol}$$

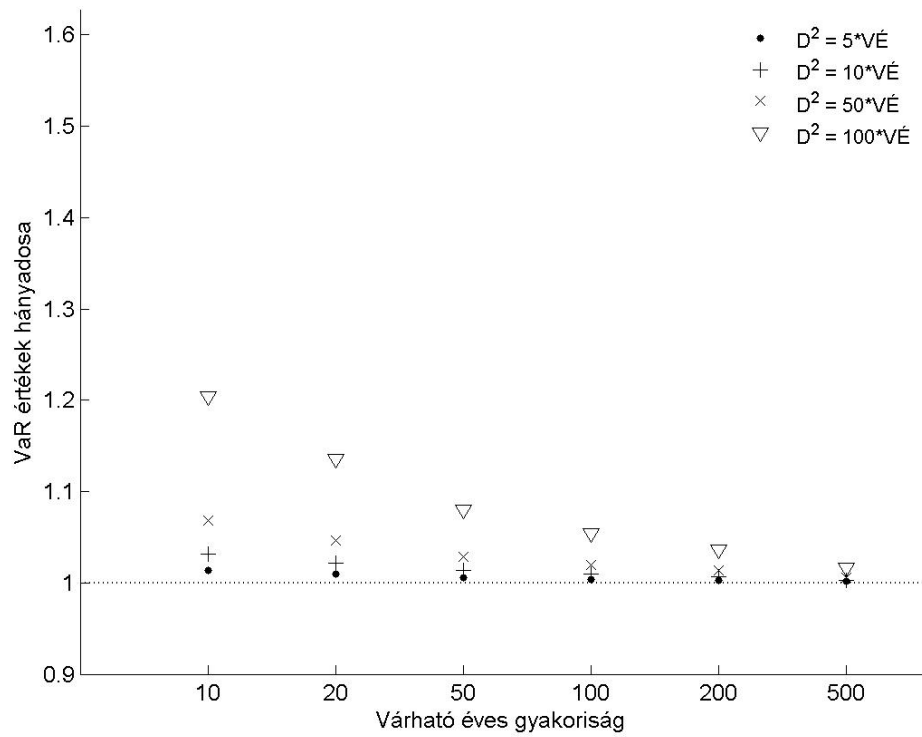
$N_{X|X \geq a}$  egy adott küszöbön túli veszteségek száma, valamint X a veszteséget, míg a egy tetszőleges valós számot jelent.

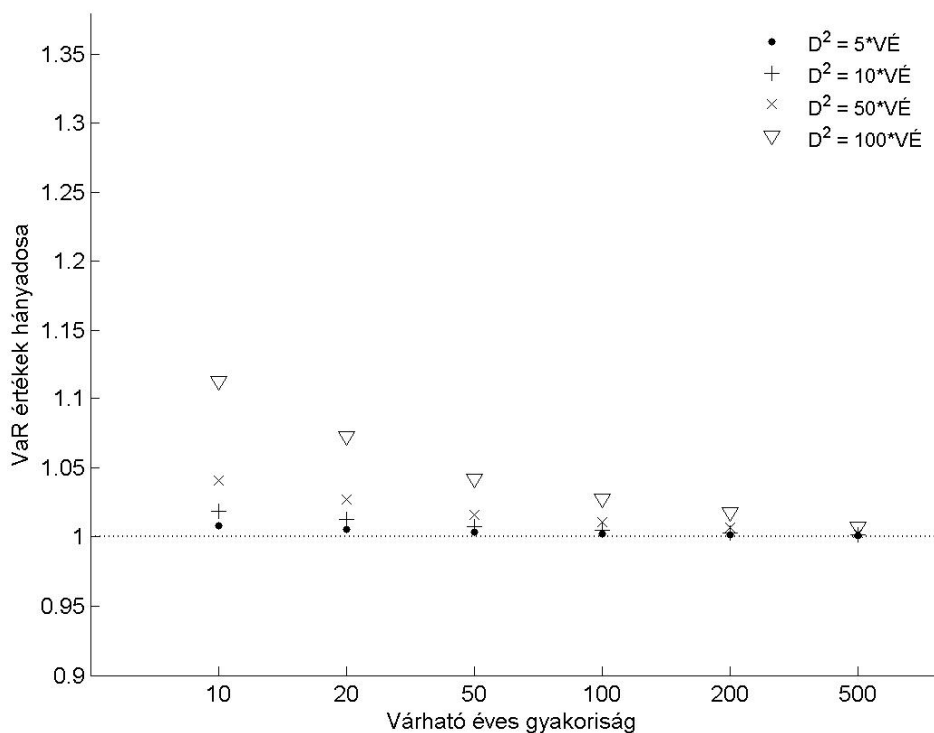
- A választott gyakoriság eloszlás korlátozott hatást gyakorol a tőkére, hiszen bizonyos körülmények (szubexponenciális veszteség eloszlás) esetén a tőkekövetelmény szempontjából csak a gyakoriság eloszlás várható értéke a lényeges, varianciája és további jellemzői nem, lásd Single Loss Approximation. Ezt az elvárást pedig az egy paraméteres Poisson eloszlás is teljesíti, tehát nem érdemes bonyolultabb modellt választani (Böcker et. al. [2005]).

A következő ábrával azt illusztráljuk, hogy milyen esetekben számít a gyakoriság eloszlás alakja annak várható értékén túl. Az ábrák azonos várható értékű Negatív Binomiális és Poisson eloszlások felhasználásával előállított tőkekövetelmények (VaR) hányadosát mutatja a gyakoriság eloszlások várható értékének a függvényében.

Három veszteségeloszlást használunk, mindegyik lognormális, de az egyik vastagabb szélű, mint a másik ( $\sigma = \{2; 2,2; 2,5\}$ ). Itt jegyezzük meg, hogy lognormális eloszlás esetén a kurtózis csak a  $\sigma$  függvénye. Minél nagyobb az eloszlás kurtózisa, annál vastagabb szél jellemző rá, lásd Balanda et. al. [1988] és Ruppert [1987].







6. ábra: az a) – c) ábrák azonos várható értékű Negatív Binomiális és Poisson eloszlások felhasználásával előállított tőkekövetelmények (VaR) hányadosát mutatja a gyakoriság eloszlások várható értékének a függvényében. A gyakoriság eloszlásokhoz tartozó lognormális eloszlások paraméterei rendre:  $\mu=10$  és  $\sigma=\{2; 2,2; 2,5\}$ .

Az ábrákról a következőket olvashatjuk le:

- A különböző gyakorisági eloszlások hatása a kis várható értékű gyakoriságok esetén fontos, a várható érték növekedésével a konkrét eloszlás fontossága csökken.
- A nagyobb varianciájú Negatív Binomiális eloszlás jelentősége annál inkább csökken, minél vastagabb szélű a veszteség eloszlás.

Megjegyezzük, hogy ez a hiba alapvetően eltörpül a veszteségeloszlás nem megfelelő választásából, valamint annak paraméterbecsléséből származó hibához képest. Ez az elsődleges oka annak, hogy a dolgozat és a szakirodalom súlypontját is sokkal inkább a veszteségeloszlás becslésével kapcsolatos problémát jelentik.

### 6.2.3. Single Loss Approximation és gyakoriság eloszlás választás

Korábban már bemutattuk, hogy a működési kockázati tőkekövetelmény, avagy az aggregált eloszlás megfelelő kvantilise szubexponenciális eloszlások esetén analitikusan is meghatározható és

$$VaR = F_s^{-1}(p) \approx F^{-1}\left(1 - \frac{1-p}{E(n)}\right). \quad (10)$$

Jól látható a 10. egyenletből, hogy az aggregált eloszlás VaR-ja csupán a gyakoriságeloszlás várható értékétől –  $E(n)$  függ. Ez összhangban van a szimulációs eredményünkkel (6. ábra), miszerint vastag szélű eloszlás esetén (pl. lognormális magas  $\sigma$  mellett) nem számít az eloszlás alakja. Mindez előtérbe helyezi a Poisson eloszlást, mint népszerű, és kedvező tulajdonságokkal rendelkező gyakoriság eloszlást, hiszen a Single Loss Approximation is azt támasztja alá, hogy elegendő csupán a gyakoriság eloszlás várható értékével foglalkozni, mivel a VaR meghatározása szempontjából a fenti modell feltételek mellett az eloszlás további jellemzői és alakja lényegtelenek.

## 6.3. A veszteségeloszlás

### 6.3.1. Gyakran alkalmazott veszteségeloszlások és illesztésük

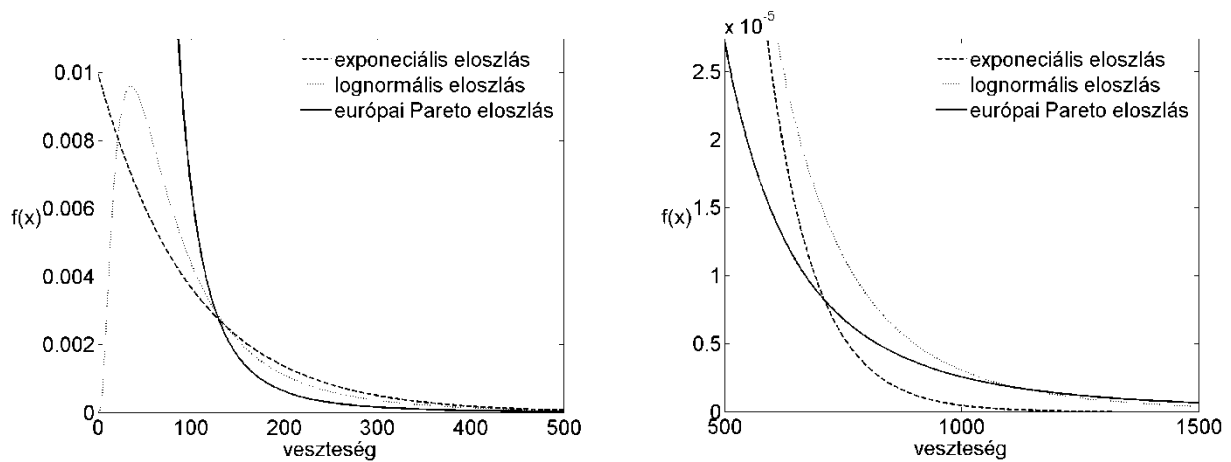
Az egyedi veszteségekről már megállapítottuk, hogy azokat nemnegatív értékű valószínűségi változóval írhatjuk le. Ennélfogva természetesen adódik, hogy nemnegatív értékű eloszláscsaládokat használjunk. További korlátozást általánosan ugyan nem tehetünk, azonban kiemelhetjük azon nevezetes eloszlások körét, melyeket a szakirodalomban is gyakorta ajánlanak. Ezeket tartalmazza a 7. táblázat. A táblázatbeli eloszlások használatának egyik alapvető oka -amellett, hogy ezek a valószínűségszámításban számos helyen használt nevezetes nemnegatív eloszlások- az, hogy egy részük korábbi pénzügyi és biztosítási területen szerzett tapasztalatok (statisztikai vizsgálatok) alapján jól illeszkedtek egyes vizsgált veszteségtípusokra. Kiemelendő, hogy ezek az eloszlások kevés paraméterrel rendelkeznek,



amely a kis mintaméretetek esetén fontos szempont a becslési megbízhatóság miatt. További alkalmas eloszlások találhatóak Panjer [2006] és Panjer et. al. [1986] művekben.

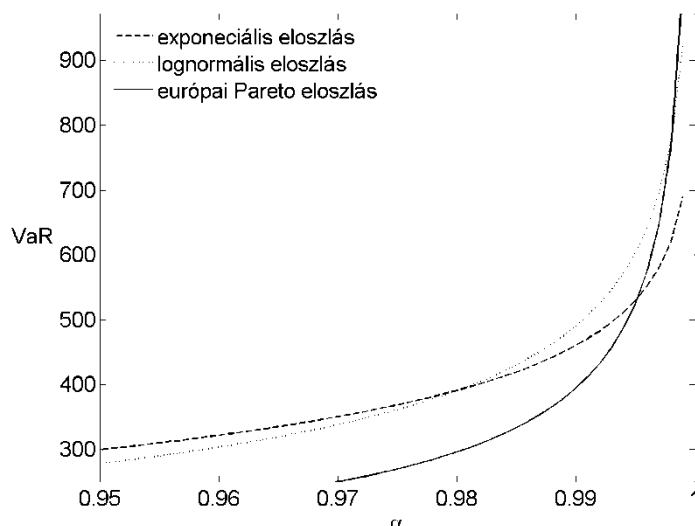
Eloszlás	Paraméterek	Sűrűségfüggvény	Megjegyzés
Exponenciális	$\lambda > 0$	$\lambda e^{-\lambda x}, x > 0$ 0, egyébként.	
Pareto (európai)	$c, a > 0$	$\frac{a}{c} \left(\frac{c}{x}\right)^{a+1}, x > c$ 0, egyébként.	Nagy veszteségek modellezése
Általánosított Pareto (GPD)	$\xi \in R, \sigma > 0$	$\frac{1}{\sigma} \left(1 + \left(\xi \frac{x - \mu}{\sigma}\right)^{-\frac{1}{\xi}}\right), x > 0$	Határeloszlás egy adott u küszöb fölött
Lognormális	$\mu \in R, \sigma^2 > 0$	$\frac{1}{\sigma x \sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, x > 0$ 0, egyébként.	Y lognormális, ha X=ln(Y) normális eloszlású
Lognormális- gamma (variance- gamma)	$\mu \in R; \sigma^2, y > 0$	$\int_0^\infty \frac{1}{x\sqrt{2\pi y}} \exp\left(-\frac{\log(x) - \mu}{2y}\right) \exp\left(-\frac{\alpha y}{\sigma^2}\right) \frac{\alpha^\alpha}{\sigma^{2\alpha}} \frac{y^{\alpha-1}}{\Gamma(\alpha)} dy$	$Z = e^{(\mu + \sigma X \sqrt{Y})}$ , ahol X=N(0,1) és $Y = \Gamma\left(\alpha, \frac{1}{\alpha}\right)$
Burr	$\alpha, \gamma, \nu > 0$	$\frac{\alpha \gamma \left(\frac{x}{\nu}\right)^\gamma}{x \left(1 + \left(\frac{x}{\nu}\right)^\gamma\right)^{\alpha+1}}$	

7. táblázat: Leggyakoribb veszteséeloszlások.



7. ábra: Az exponenciális ( $\lambda=0.01$ ), lognormális ( $\mu=4,26$   $\sigma=0,83$ ) és európai Pareto ( $\alpha= 2,41$   $c=59$ ) eloszlások sűrűségfüggvénye. Az eloszlások várható értéke és szórása egyaránt 100.

Az eloszlások paramétereinek becslése, a legjobb eloszláscsalád kiválasztása, továbbá az illeszkedés jóságának vizsgálata fontos feladat nagyobb mértékben a veszteségeloszlások, kisebb mértékben a gyakoriság eloszlások esetén. A 7. ábra jól mutatja, hogy lényegesen különböző a néhány kiemelt veszteségeloszlás sűrűségfüggvénye, különösen kiemelendő a farokeloszlások 7. ábra jobb oldalán látható eltérő jellege. Ezért nagyon fontos megfelelő eloszláscsaládot választanunk, hiszen annak a tőkekövetelmény értékére is nagy a hatása, különösen a farokeloszlások eltérő jellege miatt. Ennek szemléltetése a 8. ábra, ahol a VaR értékeket a szóban forgó egyedi eloszlásokra mutatjuk be ugyanazon paraméterértékek mellett, mint amit a 7. ábra esetében használtunk, így jól látható az eltérő jellegű sűrűségfüggvény okozta különbség. Természetesen ezeknek a hatását a teljes eloszlás VaR értékeire is bemutatjuk a későbbi szimulációs vizsgálatainkban, hiszen ezzel azt vizsgálhatjuk, hogy a rosszul választott eloszláscsalád milyen hibát okozhat a tőkekövetelményben. A paraméterbecsléshez, az eloszlások szelekciójához, az illeszkedés jóságának vizsgálatához természetesen a statisztikából jól ismert standard eljárásokat használhatjuk, például: hipotézisvizsgálatok illeszkedésre (pl.  $\chi^2$  próba, Kolmogorov-Smirnov próba, Anderson-Darling próba), grafikus eszközök (pl. hisztogram, PP-plot, QQ-plot), pontbecslési módszerek (pl. momentumok módszere, maximum likelihood módszer, kvantilis módszer). Mivel ezen eljárások alkalmazása során nem vetődnek fel működési kockázatra jellemző specifikus problémák, így ezeket jelen írásban nem tárgyaljuk részletesen.



8. ábra: Az exponenciális ( $\lambda=0.01$ ), lognormális ( $\mu=4,26$   $\sigma=0,83$ ) és európai Pareto ( $\alpha= 2,41$   $c=59$ ) eloszlások kockázatos értéke (VaR) a konfidenciaszint függvényében.<sup>21</sup>

### 6.3.2. Single Loss Approximation és az egyedi veszteségeloszlás

Korábban rávilágítottunk, hogy az aggregált eloszlás kvantilise és az egyedi veszteségeloszlás kvantilise között közvetlen összefüggés van. Az aggregált eloszlás ( $F_s$ )  $p=0,999$  konfidenciaszinthez tartozó kvantilise közelítőleg megegyezik az egyedi veszteségeloszlás

$$1 - \frac{1-p}{E(n)}$$

konfidenciaszinthez tartozó kvantilisével.

A korábban bemutatott 5. táblázat jól mutatja, hogy bizonyos esetekben mennyire magas konfidenciaszint mellett kell a VaR értéket meghatározni az egyedi veszteségeloszlásból. Ezen a ponton jogosan merül fel bennünk a kérdés, hogy mennyire pontosan tudjuk meghatározni a működési kockázati tőkekövetelményt? A táblázat 4. oszlopa az 5 év alatt megfigyelt veszteségek várható számát tartalmazza. Látható, hogy a veszteségek száma messze elmarad attól, hogy az adatokból direkt kvantilis becsléssel meghatározzuk a VaR-t. Ez azt jelenti, hogy paraméteres modellt kell illeszteni, azonban az illesztés jóságát csak

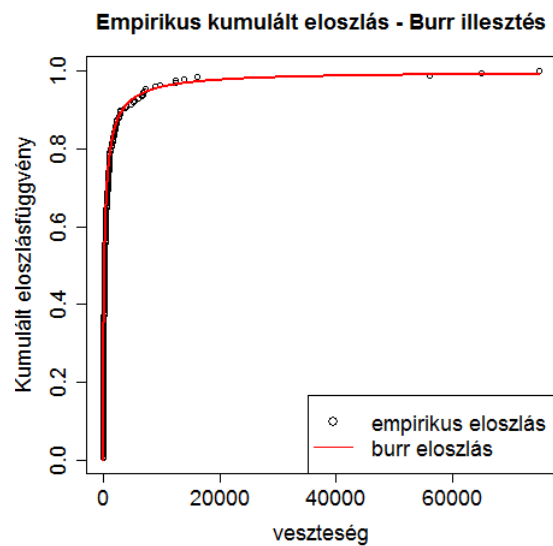
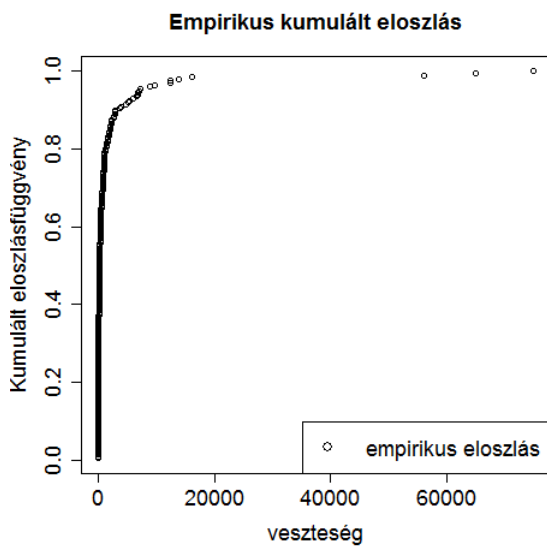
<sup>21</sup> Az eloszlások várható értéke és szórása egyaránt 100. Ez utóbbi feltételezés igen szolidnak mondható, hiszen HUNOR adatbázis statisztikáit tekintve a szórás több kategóriában is a várható érték tízszeresét közelíti<sup>21</sup>. Ugyanakkor még ezen naiv feltételezés mellett is nagy különbségek lehetnek két eloszlás VaR értékei között, miközben a konfidenciaszint maximális értéke az ábrán mindössze 0.999. A Single Loss Approximationból tudjuk, hogy a tőkéhez szükséges konfidenciaszint az éves gyakoriság várható értékétől függ, és már mérsékelt (évi 1-nél nagyobb) gyakoriságú események esetén is 0.999-nél nagyobb konfidenciaszint mellett kell az egyedi eloszlás VaR-ját vizsgálni.

abban a tartományban tudjuk tesztelni, ahol megfigyelések állnak a rendelkezésünkre, tehát a korábban említett extrapolációs problémával nézünk szembe!

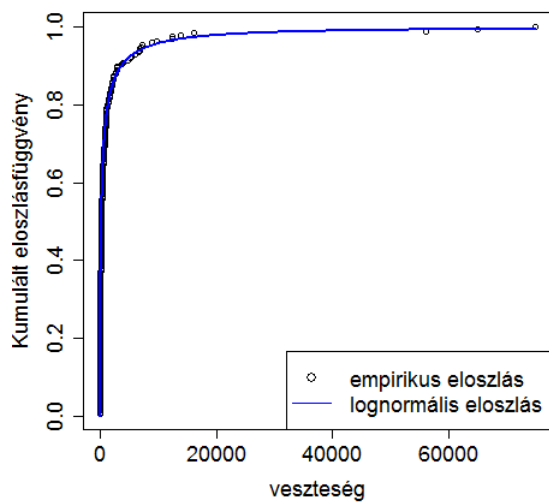
A következő példában két olyan eloszlásillesztést mutatunk meg, ahol a tesztelt eloszlások nagyon jól teljesítettek a megfigyelhető tartományban az illeszkedés vizsgálat során (lásd 8. táblázat), ugyanakkor a hozzájuk tartozó SLA tőkekövetelmények mégis nagymértékben különböznek (lásd 9. táblázat).

A 200 (évi 40) elemű mintát  $\mu=5,3$   $\sigma=2,5$  lognormális eloszlásból generáltuk, melyre Burr és lognormális eloszlásokat illesztettünk Maximum Likelihood módszerrel.

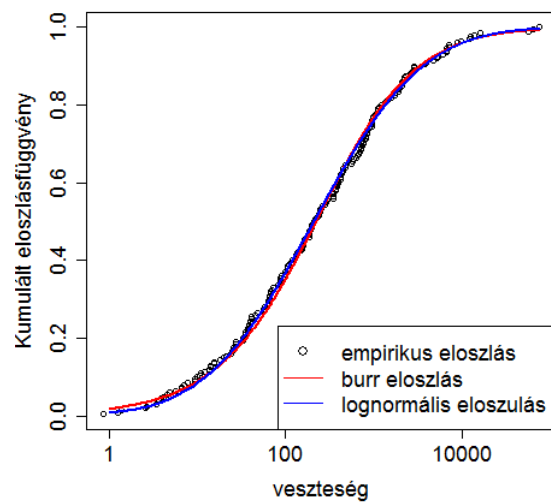
Számos hasonló példa található a szakirodalomban, lásd pl. Mignola et. al. [2006] cikke.



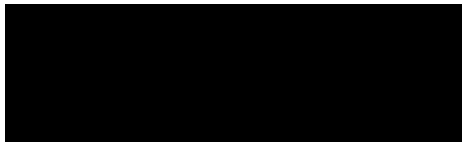
Empirikus kumulált eloszlás - lognormális illesztés



Eloszlások illesztése



9. ábra: 200 elemű minta lognormális  $\mu=5.3$   $\sigma=2.5$  paraméterekkel.<sup>22</sup>



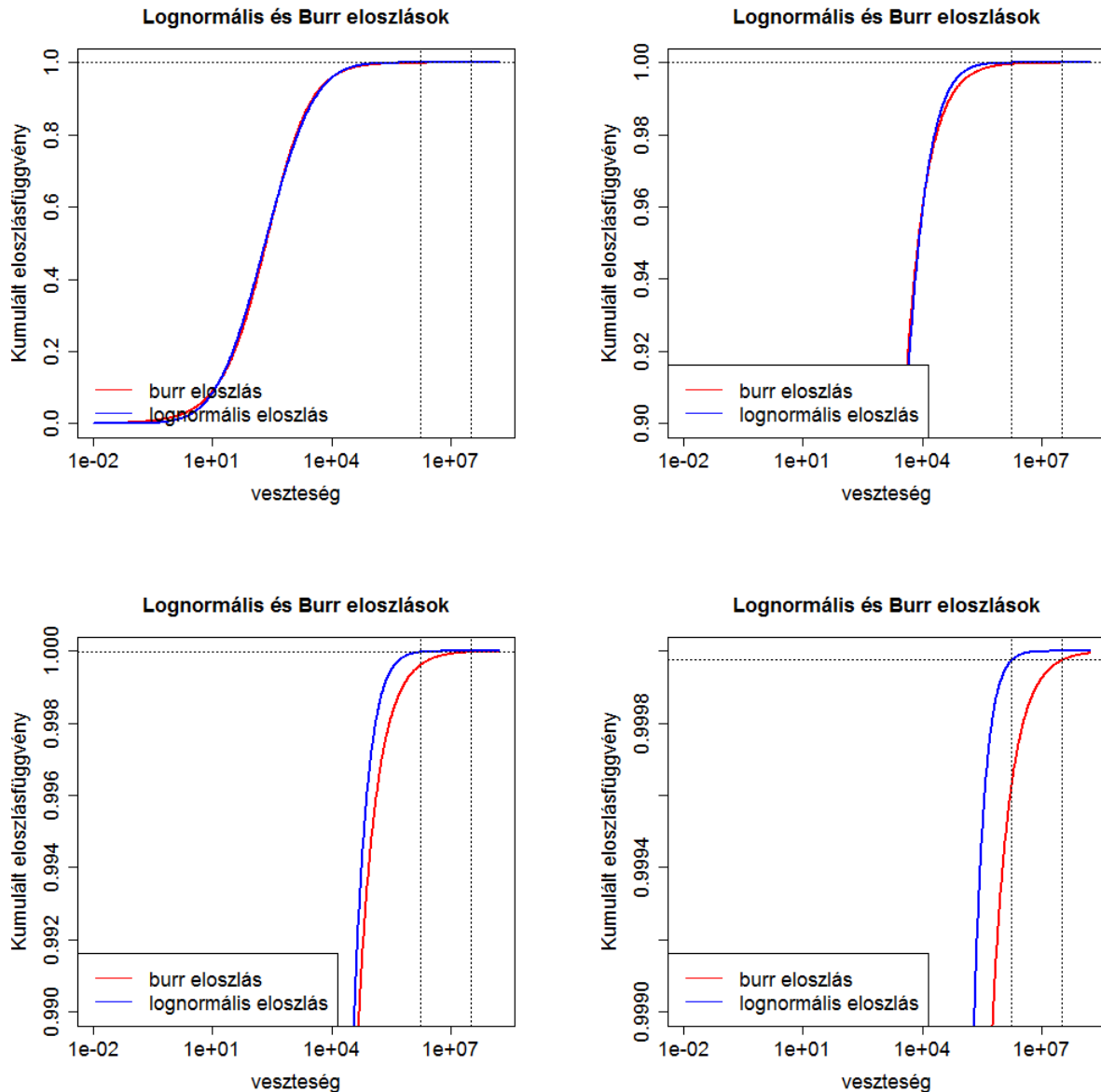
8. táblázat: Illesztés jósága (p-values) Kolmogorov-Smirnov (KS) és Anderson-Darling (AD) tesztek alapján. A döntés nem egyértelmű, hiszen a KS teszt alapján a Burr, míg az Anderson-Darling teszt alapján a lognormális eloszlás a preferált!

<sup>22</sup> a) ábra: Kumulált empirikus eloszlás.

b) ábra: Burr eloszlás illesztése a kumulált empirikus eloszlásra.

c) ábra: Lognormális eloszlás illesztése a kumulált empirikus eloszlásra.

d) ábra: Lognormális eloszlás és Burr eloszlás illesztése a kumulált empirikus eloszlásra. Az x tengely szerinti értékek (vesztesség) logaritmusát vettük, miközben az x tengely a logaritmus nélküli értéket mutatja.



10. ábra: A 200 elemű lognormális ( $\mu=5,3$ ,  $\sigma=2,5$ ) mintára illesztett Burr és lognormális kumulált eloszlások logaritmikus x tengelyen. Az a)-d) ábrák egyre inkább csak az eloszlás szélére fókuszálnak, míg a függőleges szürke vonalak a Single Loss Approximation alapján kalkulált VaR értékeket jelölik.

p	igazi VaR	Burr VaR	Lognormális VaR
0,9	4 934	3 310	3 643
0,99	67 230	48 797	37 202
0,999	453 877	599 296	203 410
0,9999	2 185 907	7 159 036	823 549
0,999975	5 071 096	31 789 076	1 741 037

9. táblázat: Az illesztett eloszlások VaR értékei különböző konfidencia szintek mellett. A táblázat utolsó rekordjából jól látható, hogy a két kiválóan illeszkedő eloszlás teljesen különböző tőkekövetelményekhez vezet, miközben mindkettő óriási hibával becsli a tényleges tőkét.

A két különböző eloszlás, miközben a megfigyelhető tartományban mindkettő nagyon jól illeszkedett, nagymértékben eltérő tőkekövetelményt eredményezett. Nagyjából 20-szoros a különbség a tőkekövetelmény értékek között! Hangsúlyozzuk, hogy ezen szimulált körülmények között nem kellett foglalkoznunk semmilyen egyéb más becslést befolyásoló tényezővel, azok hatása tovább rontaná a helyzetet.

#### 6.4. Adatgyűjtési küszöb

##### 6.4.1. Adatgyűjtési küszöb hatása a tőkére

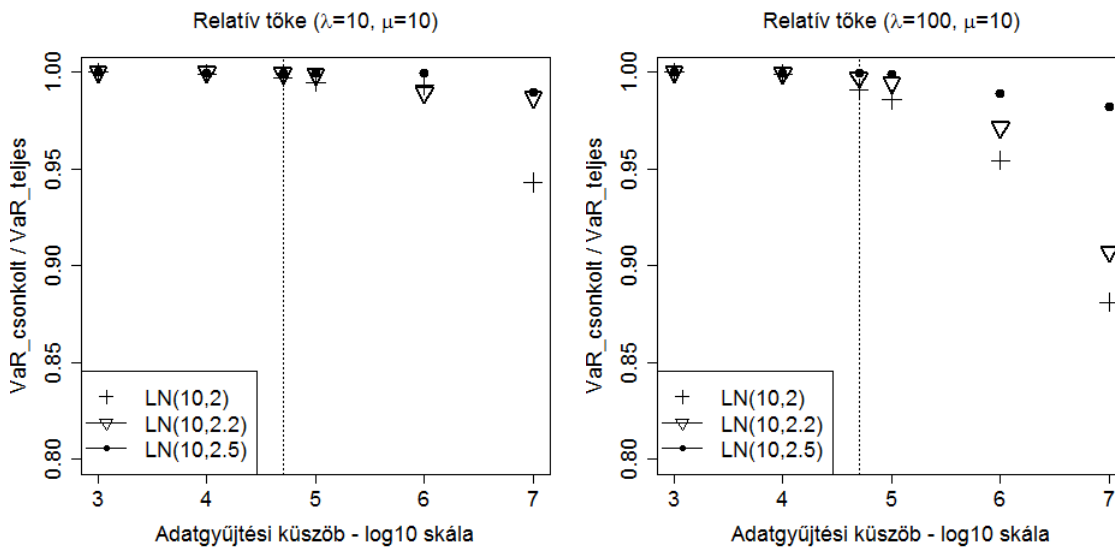
Fontos ismernünk az adott pénzügyintézet adatrögzítési gyakorlatát. Az intézetekben ugyanis csak egy adott szint fölötti veszteségek kerülnek be a működési kockázatok adatbázisába, így a felhasználható mintába is. Ez esetben tehát egy szűrt mintánk van, a szűrési feltételt a megadott limit jelenti. Ekkor, ha körültekintően akarunk eljárni, akkor az egyedi veszteségeknél és gyakoriságoknál ezen szűrési feltételt figyelembe véve kell elvégezni a paraméterbecsléseket (akár momentumok módszerének alkalmazása esetén, akár maximum likelihood becslés esetén). Ezt a korrekciót a maximum likelihood módszerre meg is mutatjuk, az érdeklődő olvasónak javasoljuk Klugman et. al. [2004] és Panjer [2006] műveit.

A HUNOR adatkonzorcium tagjai az 50 000 forint feletti veszteségeket kötelesek továbbítani az HUNOR adatbázisba. Felmerül a kérdés, hogy tőkekövetelmény szempontjából van-e jelentősége az 50 000 forint alatti veszteségeknek? Itt most nem foglalkozunk azzal hogy milyen következménye van a becslésre nézve annak, hogy eldobjuk a minta egy részét, ezt a következő pontban tárgyaljuk. A fenti kérdés kapcsán Mignola et. al. [2006] ötlete alapján a következő Monte Carlo szimulációs kísérletet végeztük el.

1. Poisson folyamatból  $\lambda$  éves gyakorisággal lognormális egyedi veszteségeket generáltunk.
2. Az előre rögzített  $K$  küszöb fölötti veszteségekből meghatároztuk az éves aggregált veszteséget.
3. Az 1-2 lépéseket megismételtük 10 000-szer, és meghatároztuk az aggregált veszteségek eloszlásának  $p=0,999$  konfidenciaszint melletti  $VaR_K$  értékét.

4. Megképeztük a  $\frac{VaR_K}{VaR_0}$  hányadost, ahol  $VaR_K$  a  $K=k$ , míg  $VaR_0$  a  $K=0$  értékhez tartozó kockázatotott érték.

A szimulációs eredményeket a 11. ábra tartalmazza.



11. ábra: Tőke relatív alakulása az adatgyűjtési küszöb függvényében.<sup>23</sup>

A fenti ábrákból két dolgot olvashatunk le:

- minél kevésbé vastag szélű az eloszlás, annál inkább számít az adatgyűjtési küszöb (mindkét ábrán egyaránt látható, hogy kisebb  $\sigma$  mellett romlik a becslés);
- minél nagyobb az események éves gyakorisága, annál inkább számít az adatgyűjtési küszöb (eloszlások páronkénti összehasonlítása mutatja, hogy ugyanazon küszöb mellett nagyobb gyakoriság esetén jobban alábecsüljük a tőkét).

<sup>23</sup> Az adatokat Poisson ( $\lambda=[10; 100]$ ) gyakoriság eloszlásból és  $\mu=10$  paraméterű lognormális eloszlásból szimuláltuk. A lognormális eloszlás  $\sigma$  paraméterét változtattuk ( $\sigma=[2; 2,2; 2,5]$ ), lásd ábra. Az y tengely az adatgyűjtési küszöb fölötti veszteségekből és az összes veszteségből számolt tőke (VaR) arányát mutatja. Az x tengelyen az adatgyűjtési küszöb látható logaritmus skálán. Az függőleges szaggatott vonal a Magyarországon tipikus 50 000 forintos adatgyűjtési küszöböt jelenti.



Láthatjuk, hogy önmagában a kis veszteségek eldobása tipikusan nem jelent komoly veszélyt a tőke meghatározására nézve, habár erre számítottunk a Single Loss Approximation és az empirikus tanulmányok pl. Aue et. al. [2006] alapján is. A problémát inkább az jelenheti, hogy romlik az egyedi veszteségeloszlás becslése a csökkenő mintaelemszám miatt, ami a tőke növekvő instabilitásához vezet. Ugyanakkor az továbbra is kérdés az extrapolációs probléma ismeretében, hogy mennyire kell az egyedi veszteség eloszlás becslésének jóságával foglalkoznunk, hiszen a GoF tesztek úgyszólván csak a megfigyelhető tartományra korlátozódnak, és a becslött tőke egy jó illesztés esetén is instabil lehet, lásd pl. 8. táblázat és 9. táblázat eredményei.

#### 6.4.2. Adatgyűjtési küszöb és az egyedi veszteségeloszlás paramétereinek Maximum Likelihood becslése

Az intézmények adatgyűjtési küszöbe ugyan közvetlenül kis hatással van a tőkére, ám jelentős hatással bírhat az egyedi veszteségeloszlás paramétereinek a becslésére. A legnépszerűbb paraméterbecslési módszerek a következők:

- momentumok módszere (moment matching),
- súlyozott momentumok módszere (probability weighted moments),
- kvantilisek módszere (quantile matching),
- maximum likelihood módszer.

A maximum likelihood becslés népszerűsége az alábbi tulajdonságokkal magyarázható:

- konzisztens,
- hatásos,
- aszimptotikus normalitása a paramétereknek,
- invariáns.

A fentiekén túl a maximum likelihood módszer lehetővé teszi, hogy konfidencia intervallumot adjunk meg a becslésre, a Hesse mártix felhasználásával következtetést vonhatunk le a paraméterek korreláltságáról.

Ha a maximum likelihood becslés létezik, akkor a fenti tulajdonságok garantáltak, és ez a módszer legnagyobb előnye! Az érdeklődő olvasó többet tudhat meg a maximum likelihood becslésről Rychlik et. al. [2006] könyvéből.

Tetszőleges  $K > 0$  adatgyűjtési küszöb esetén a paraméterek korrekt becslése a feltételes eloszlásból ( $f_K$ ) végezhető el, hiszen azt tudjuk megfigyelni, a  $K$  küszöb feletti minta áll rendelkezésre. Az  $f_K$  nem negatív eloszlás balról csonkolt, és a következő összefüggésben áll a nem feltételes  $f$  eloszlással:

$$f_K(x) = f(x|K) = \frac{f(x)}{1 - F(K)} \quad (11)$$

Az  $f$  eloszlás log-likelihood függvénye legyen  $l_n = \sum_{i=1}^n \ln(f(x_i; \Theta))$ , ahol  $\Theta$  az  $f$  paramétereinek a vektora. Az  $f_K$  feltételes eloszlás log-likelihood függvénye pedig:

$$l_n^K(\Theta) = \sum_{i=1}^n \ln(f(x_i; \Theta)) - n \times \ln(1 - F(K; \Theta)). \quad (12)$$

Lognormális eloszlást feltételezve a log-likelihood függvény a következő alakú:

$$l_n^K(\mu, \sigma) = -n \times \ln(\sqrt{2\pi\sigma^2}) - \sum_{i=1}^n \ln(x_i) - \frac{1}{2} \sum_{i=1}^n \frac{(\ln(x_i) - \mu)^2}{\sigma^2} - n \times \ln(1 - F(K; \mu, \sigma)). \quad (13)$$

Problémánkat az  $-n \times \ln(1 - F(K; \mu, \sigma))$  tag okozza. A maximum likelihood becslés ennyivel több feltételes eloszlás becslése esetén a feltétel nélküli esethez képest. Feltétel nélküli esetben a becslött paraméterek aszimptotikusan normális eloszlásúak és korrelálatlanok, amit a az MLE becslés információs mátrixának inverze  $[I(\Theta)]^{-1}$  ad meg számunkra<sup>24</sup> (Lynch [2007]).

$$[I(\mu, \sigma)]^{-1} = \begin{bmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{2\sigma^4}{n} \end{bmatrix}.$$

<sup>24</sup> Az információs mátrixot a következőképpen definiáljuk:  
, ahol  $H(\Theta)$  a log-likelihood becslés Hesse mátrixa.

Ez, feltételes eloszlás esetében megváltozik, és a Cramer Rao féle korlát lognormális eloszlás esetén a küszöb függvényében a következőképpen alakul (Roehr [2002]):

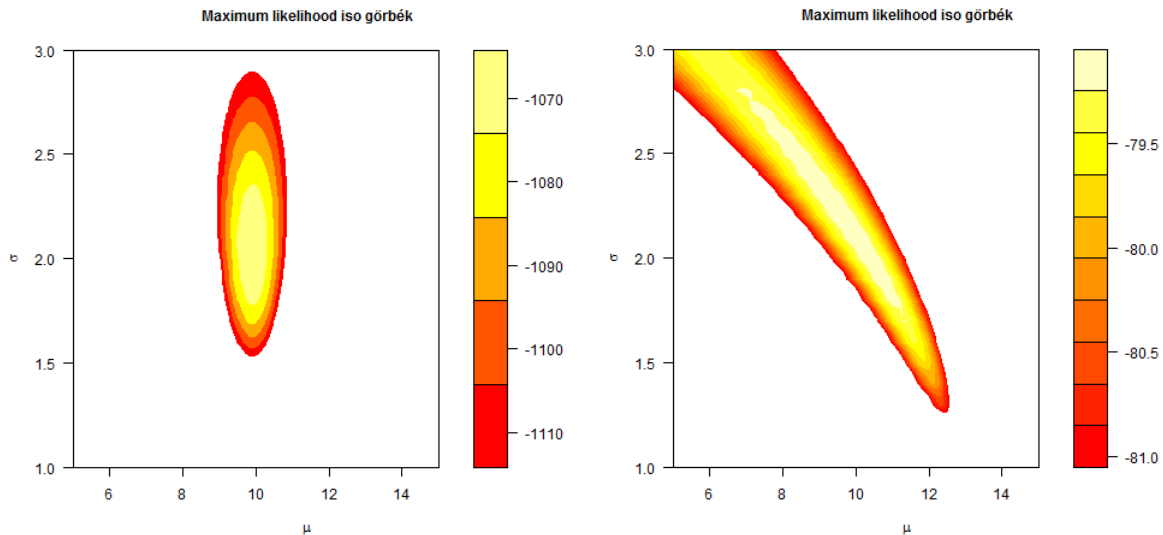
$$[I(z)]^{-1} = \frac{\sigma^2}{2 + H(H-u)(u(H-u)-3)} \begin{bmatrix} 2 + Hu(u(H-u)-1) & H(u(H-u)-1) \\ H(u(H-u)-1) & 1 - H(H-u) \end{bmatrix}, \quad (14)$$

ahol  $H = H(u) = \frac{h(u)}{\Phi(u)}$ .

Továbbá  $h(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$  és  $u(x) = \frac{\log(x) - \mu}{\sigma}$ ,

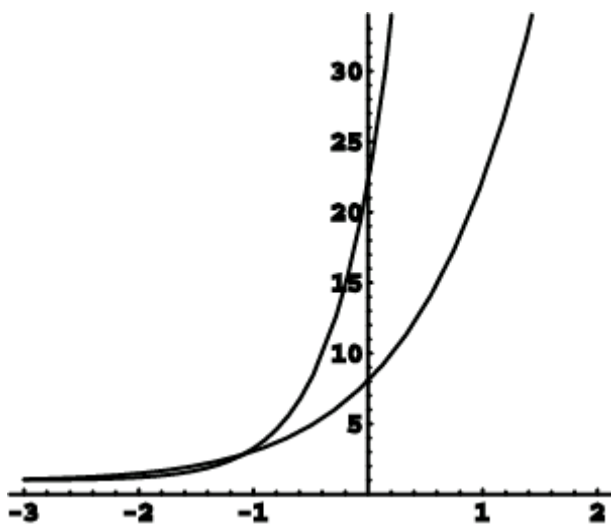
míg  $\bar{\Phi} = 1 - \Phi$ , ahol  $\Phi$  a standard normális eloszlás eloszlásfüggvénye.

Azonnal látható, hogy a hogy a vegyes tagok megjelennek az információs mátrix inverzében, tehát a becült paraméterek korrelálttá válnak, illetve a Cramer Rao határ a küszöbnek (u-nak, illetve közvetve x-nek) a függvénye. Az előbbi következménye, hogy csonkolt (feltételes) eloszlás használata esetén a log-likelihood függvény iso felülete eltorzul, megnyúlik, keskeny formájúvá válik (lásd 12. ábra), emiatt a becslés során instabil helyzetet eredményez (Frachot et. al. [2004]. Ennek következménye, hogy a konfidencia intervallumok kitágulnak, információértéküket elveszítik. Ettől kezdve kérdésessé válik, mennyire bízhatunk, vagy bízhatunk-e egyáltalán a becült paraméterekben? Léteznek javaslatok, robosztus becslések a probléma bizonyos szintű kezelésére, ilyen pl. Cope [2011] munkája, ugyanakkor itt nem célunk ennek bemutatása, elsődlegesen a probléma természetére kívánjuk felhívni a figyelmet.



12. ábra: Lognormális eloszlásból ( $\mu=10$  és  $\sigma=2$ ) generált véletlen számokra illesztett lognormális és feltételes lognormális eloszlás log-likelihood függvényének iso felülete. Mindkét esetben 500 elemű mintát használtunk fel a becsléshez, ugyanakkor a 2. esetben a minta csak 200 000 forintnál nagyobb veszteségeket tartalmazott.

Az adatgyűjtési küszöb negatív hatását a becslésre Roehr [2002] is megállapítja, miközben számos releváns eloszlásra levezeti a Cramer-Rao határ (Cramer rao bound) alakulását az adatgyűjtési küszöb függvényében. A 13. ábra lognormális eloszlás esetében mutatja a feltételes és feltétel nélküli eloszlás Cramer Rao féle határának hányadosát a „küszöb” –  $u$  függvényében. Ugyanezekre a következtetésre jut Douglas et. al. [2012] cikkében miközben számos eloszlásra vizsgálta a csonkolás következményeit.



13. ábra: A feltételes és feltétel nélküli lognormális eloszlás Cramer Rao féle határának hányadosa a „küszöb” –  $u$  függvényében  $\mu$  és  $\sigma$  paraméterekre egyaránt. Forrás: Roehr [2002]

## 6.5. Azonos mintaelemek

A működési kockázati veszteségek között számos esetben találkozhatunk azonos értékekkel. Ezek az egyes veszteséggeneráló folyamatok technikai jellegéből adódnak, melynek legekleatásabb esetét a kártyaveszteségek jelentik. A hétköznapi ügyfelek pl. tipikusan 50 000, 100 000 vagy 200 000 forintot állítanak be készpénzfelvételi limitnek, vagy sokszor a bankok is ezek valamelyikét vagy hasonló értéket adnak meg kártyakibocsátáskor, melyen az ügyfél nem változtat. Ebből adódóan egy kártyavisszaélés esetén, ahol csak a kártya, vagy napi készpénzfelvételi limit szab gátat az eltulajdonított összegnek ezek az értékek jelennek működési veszteséggént. Emiatt ebben és az ehhez hasonló kockázati osztályokban egy nevezetes eloszlás sem illeszkedik igazán jól a mintára. A becslés során ezzel a problémával foglalkozni kell, a modellező számára az eloszlások (folytonos és diszkrét) keverése, az azonos mintaelemek szétkenése a környező intervallumban jelenthet megoldást.

## 7. Tőkekövetelmény-számítás gyakorlati megvalósítása

### 7.1. Az összetett eloszlás (LDA) és a tőkekövetelmény

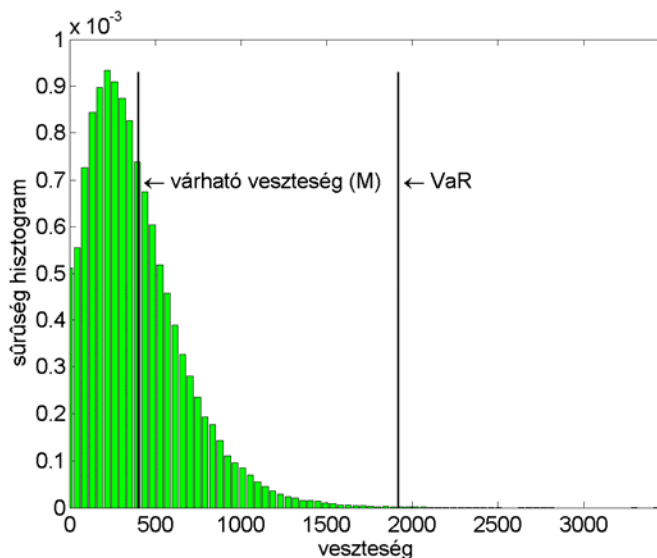
A veszteség- és a gyakoriságeloszlások ismeretében következő feladatunk a teljes eloszlás illetve a teljes eloszlás kvantiliseinek meghatározása vagy becslése. Habár eddig is alkalmaztuk ezeket a módszereket, hiszen az elvégzett kísérletek eredményei (VaR – tőke értékei) ezt megkövetelték, ugyanakkor annak számítási módját és az ahhoz közvetlenül kapcsolódó számítási problémákat nem részleteztük. A teljes / aggregált eloszlás meghatározására több módszer is rendelkezésünkre áll a Loss Distribution Approach keretein belül, ezeket az alábbiakban tételesen végigvesszük. A továbbiakban  $X$  egy olyan valószínűségi változót jelöl, amelynek eloszlása megegyezik az egyedi veszteségeloszlással. Emlékeztetőül hangsúlyozzuk, hogy az egyedi veszteségek azonos eloszlásúak a feltételezésünk szerint egy adott veszteségosztályban.

#### 7.1.1. Monte Carlo szimuláció

A leginkább kézenfekvő és leginkább javasolt módszer a felmerülő matematikai nehézségek miatt a Monte Carlo módszer, melynek segítségével közvetlenül adhatunk becslést adott biztonsági szint mellett a keresett kvantilisre, azaz a Value at Risk értékre.

Ehhez mindössze egy nagy méretű (azaz nagy szimulációs számú) mintát kell generálnunk a teljes veszteségeloszlásra (nyilvánvalóan gyakoriságok és veszteségadatok generálásával), melyhez elegendő ismernünk az egyedi veszteség- és gyakoriságeloszlásokat. Majd a generált mintából számíthatjuk ki a keresett kvantilist, amely a tőkekövetelményünk becslése lesz. Algoritmus a következő:

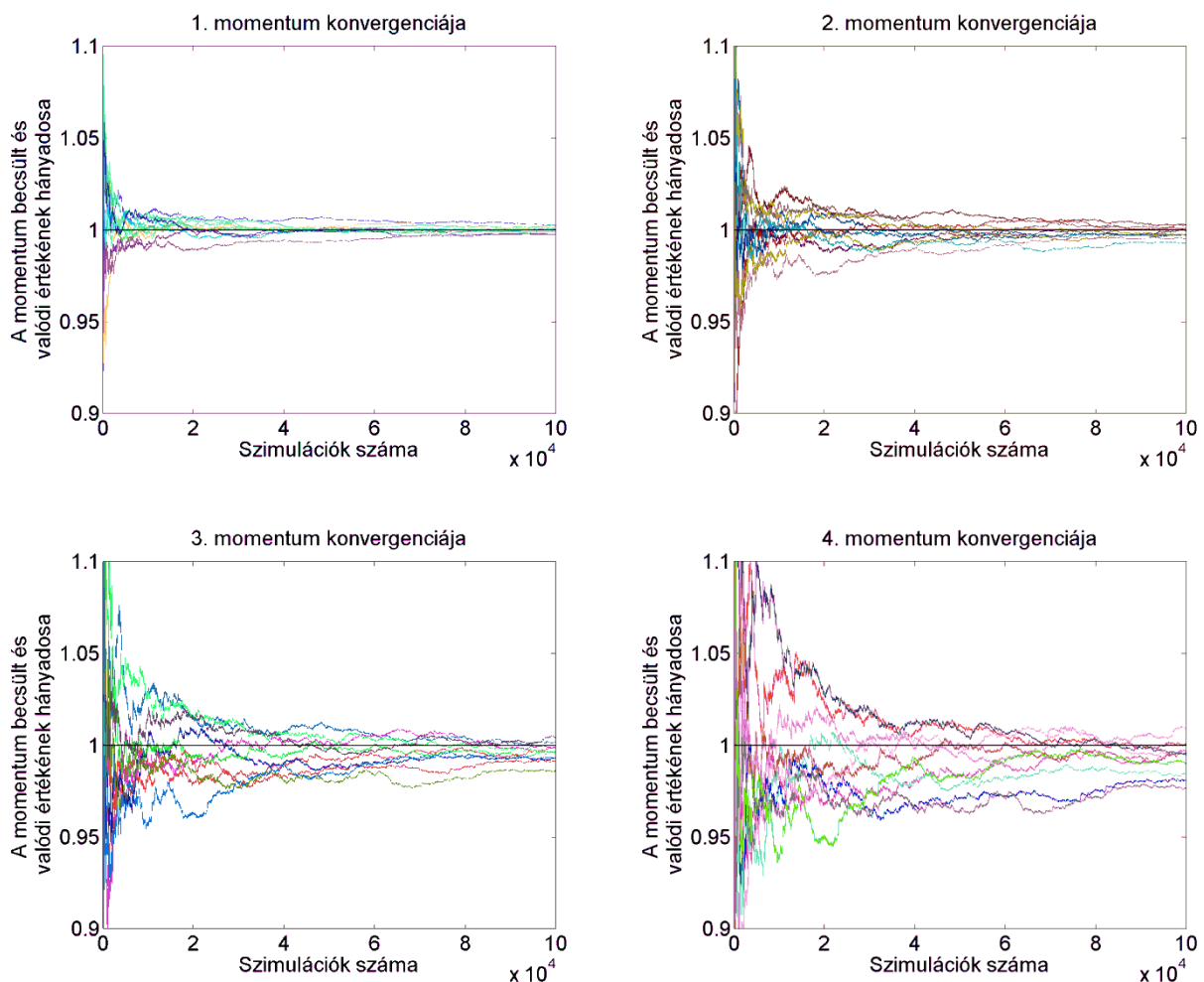
- 1) Generáljunk egy véletlen számot ( $n_i$ ) az eseményszám eloszlásból;
- 2) Generáljunk  $n_i$  db véletlen számot ( $X_1, \dots, X_{n_i}$ ) az egyedi veszteség eloszlásból;
- 3) Képezzük az  $S_i = X_1 + \dots + X_{n_i}$  összeget;
- 4) 1-3 lépéseket hajtsuk végre  $R$ -szer, ahol  $R$  egy kellően nagy pozitív valós szám;
- 5) Az  $R$  elemű mintából ( $S_1, \dots, S_R$ ) határozzuk meg az  $S$  eloszlás jellemzőit, köztük annak VaR-ját;



14. ábra: A várható veszteség és a kockázatos érték, mint statisztikai jellemzők illusztrálása a Monte Carlo szimulációval előállított Poisson ( $\lambda=4$ ) - lognormális ( $\mu=4,26$   $\sigma=0,83$ ) (P-L) összetett eloszláson.

Az eljárás használatakor kritikus a szimulációs szám nagysága. A 15. ábra és a 10. táblázat jól mutatja, hogy a kis szimulációs szám (pl. néhány ezer) már érezhető becslési hibát eredményezhet. Lognormális egyedi veszteségeloszlást feltételezve a 15. ábra mutatja, hogy mennyire gyors (vagy éppen lassú) a teljes veszteségeloszlás tapasztalati momentumainak a konvergenciája, ahol a becslült értékeket a valódi értékkel vett hányadosában mutatjuk be. Ez

természetesen akkor érdekes, ha a paramétereket a momentumokból becsüljük. A 10. táblázat pedig már a hányadosok szórását és a megfelelő VaR becslések átlagát és szórását (standard hibáját) is mutatja. Különösen kiemelendő, hogy a táblázatban a szimulációk alacsony száma esetén a VaR standard hibája már jelentős lesz, annál magasabb természetesen, minél magasabb a biztonsági szint. Kiemelendő a szabályzó által megkövetelt 99,9% konfidenciaszint mellett nagymértékű a standard hiba (VaR999 SE), ugyanakkor ez a szimulációk számával ellensúlyozható. A MC becslést minden szimulációs szám mellett 100-szor ismételtük meg.



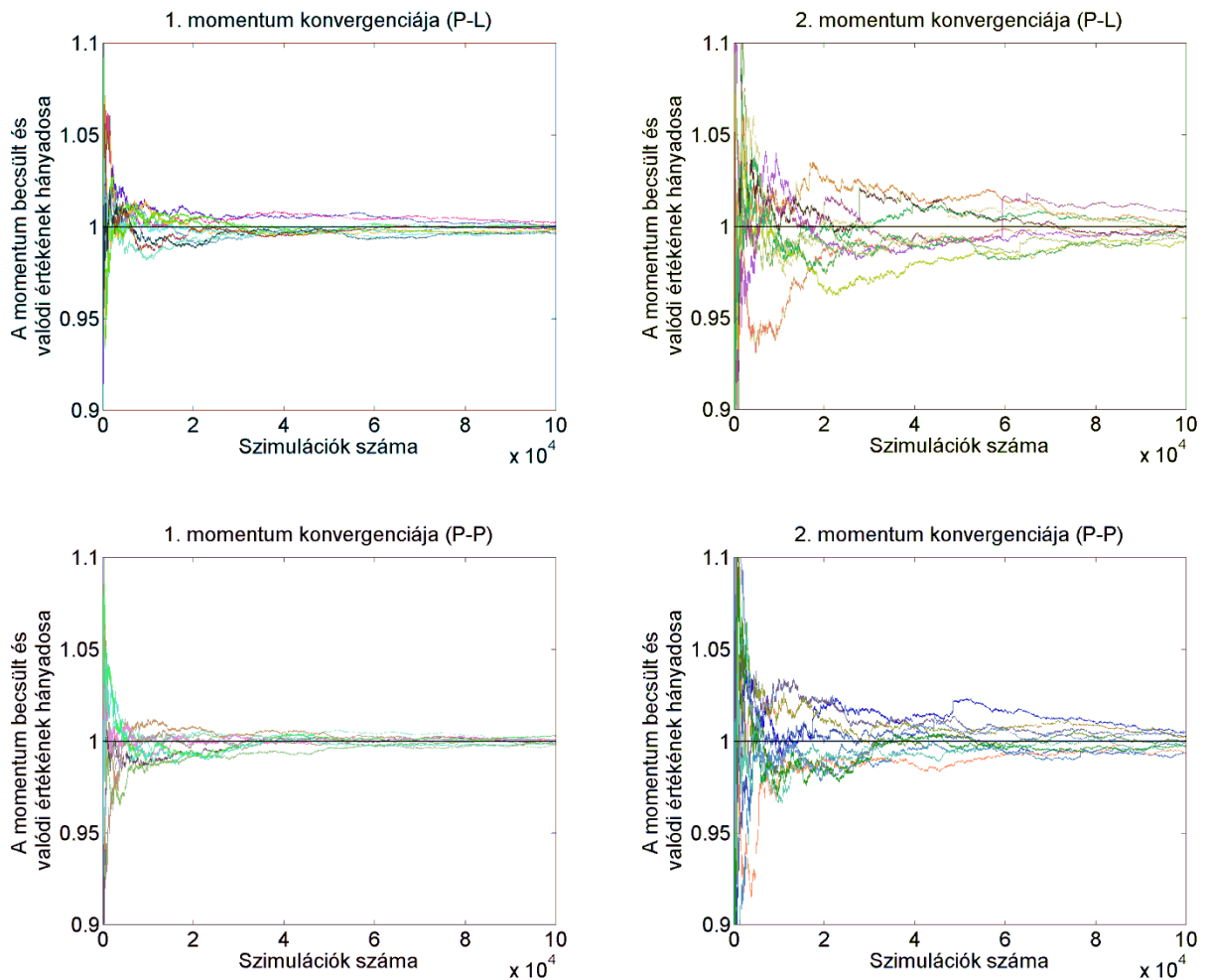
15. ábra: Összetett Poisson ( $\lambda=4$ ) eloszlás első négy momentumának a konvergenciája a szimulációs szám függvényében. A veszteségeloszlás lognormális eloszlású  $\mu=4,26$   $\sigma=0,83$  paraméterekkel.

Összetett eloszlás meghatározásához használt szimulációk száma	1000	5000	10 000	50 000	100 000
M1 becsült/igazi átlag	1,0009	0,9999	0,9997	1,0001	1,0000
M2 becsült/igazi átlag	1,0031	1,0001	0,9997	0,9999	0,9998
M3 becsült/igazi átlag	1,0096	1,0008	1,0002	0,9992	1,0000
M4 becsült/igazi átlag	1,0305	1,0000	0,9996	0,9971	1,0060
M1 becsült/igazi SE	0,0217	0,0098	0,0071	0,0030	0,0025
M2 becsült/igazi SE	0,0504	0,0227	0,0168	0,0070	0,0058
M3 becsült/igazi SE	0,1170	0,0517	0,0363	0,0158	0,0139
M4 becsült/igazi SE	0,3596	0,1359	0,0862	0,0522	0,0663
VaR95 átlag	937	932	932	932	932
VaR99 átlag	1 313	1 311	1 310	1 309	1 309
VaR999 átlag	1 978	1 926	1 914	1 901	1 902
VaR95 SE	35	15	10	5	4
VaR99 SE	74	36	27	11	8
VaR999 SE	342	133	95	38	29
VaR95 max - VaR95 min	177	77	52	21	20
VaR99 max - VaR99 min	382	158	129	52	36
VaR999 max - VaR 999 min	1 854	584	510	185	151

10. táblázat: A momentumok és a VaR jellemzőinek alakulása adott szimulációs szám mellett, Poisson ( $\lambda=4$ ) – lognormális ( $\mu=4,26$   $\sigma=0,83$ ) összetett eloszlás esetén. A VaR becslést 100-szor végeztünk el, M1 – M4 az összetett eloszlás első négy momentumát jelölik.

Frachot et. al. [2001] ötlete alapján egy összehasonlítást is végeztünk a konvergencia sebességekben lognormális és Pareto veszteségeloszlások esetén, amelyeknek azonos az első kettő momentuma. Az eredményeket a 16. ábra valamint a 11. táblázat és 12. táblázat adatai tartalmazzák. A konvergencia sebessége a két eloszlás esetében (az adott feltételek mellett) közel azonosnak bizonyult.





16. ábra: Poisson ( $\lambda=4$ ) lognormális ( $\mu=4,11$   $\sigma=1$ ) (P-L) és Poisson ( $\lambda=4$ ) amerikai Pareto ( $\alpha=4,9$   $\beta=390$ ) (P-P) összetett eloszlások első két momentumának a konvergenciája. Mindkét veszteségeloszlás esetében a várható

érték 100, míg a szórás 130. Az amerikai Pareto eloszlás eloszlásfüggvénye:  $1 - \left(\frac{\beta}{\beta + x}\right)^\alpha$ , ha  $x > 0$ , egyébként 0.

Összetett eloszláshoz használt szimulációk száma	1000	5000	10 000	50 000	100 000
M1 becsült/igazi átlag	0,9962	0,9987	0,9994	0,9996	0,9998
M2 becsült/igazi átlag	0,9965	0,9949	0,9976	0,9995	1,0000
M1 becsült/igazi SE	0,0252	0,0109	0,0073	0,0039	0,0026
M2 becsült/igazi SE	0,0675	0,0302	0,0204	0,0103	0,0072
VaR95 átlag	1 006	1 009	1 009	1 010	1 010
VaR99 átlag	1 538	1 527	1 530	1 530	1 531
VaR999 átlag	2 680	2 464	2 490	2 504	2 506
VaR95 SE	41	18	13	5	4
VaR99 SE	112	44	28	16	11
VaR999 SE	624	250	161	78	53
VaR95 max - VaR95 min	184	97	70	25	18
VaR99 max - VaR99 min	518	226	144	84	52
VaR999 max - VaR 999 min	2 895	1 361	920	435	280

11. táblázat: A momentumok és a VaR jellemzőinek alakulása adott szimulációs szám mellett, Poisson ( $\lambda=4$ ) – lognormális ( $\mu=4,11$   $\sigma=1$ ) összetett eloszlás esetén. A VaR becslést 100-szor végeztünk el, M1 – M4 az összetett eloszlás első négy momentumát jelölik. A lognormális eloszlás várható értéke 100, szórása 130.

Összetett eloszlás meghatározásához használt szimulációk száma	1000	5000	10 000	50 000	100 000
M1 becsült/igazi átlag	0,9955	0,9998	0,9995	1,0006	1,0004
M2 becsült/igazi átlag	0,9860	0,9986	0,9996	1,0013	1,0006
M1 becsült/igazi SE	0,0241	0,0125	0,0082	0,0038	0,0027
M2 becsült/igazi SE	0,0528	0,0312	0,0212	0,0089	0,0066
VaR95 átlag	1 010	1 021	1 021	1 021	1 021
VaR99 átlag	1 484	1 499	1 502	1 504	1 503
VaR999 átlag	2 399	2 318	2 333	2 331	2 331
VaR95 SE	36	19	13	6	4
VaR99 SE	80	46	34	14	10
VaR999 SE	415	204	126	56	41
VaR95 max - VaR95 min	168	103	86	28	18
VaR99 max - VaR99 min	421	259	164	74	48
VaR999 max - VaR 999 min	2 223	1 224	645	279	207

12. táblázat: A momentumok és a VaR jellemzőinek alakulása adott szimulációs szám mellett, Poisson ( $\lambda=4$ ) – Pareto ( $\alpha=4,9$   $\beta=390$ ) összetett eloszlás esetén. A VaR becslést 100-szor végeztünk el, M1 – M4 az összetett eloszlás első négy momentumát jelölik. A Pareto eloszlás várható értéke 100, szórása 130.

### 7.1.2. Panjer rekurzió

Egy másik lehetőség a teljes eloszlás egzakt meghatározása (az 'összetevők' ismeretében, s nem csak becslése), majd abból a kvantilis kiszámolása. A teljes eloszlást azonban nem egyszerű meghatározni. Az egyik legismertebb eszköz ehhez a Panjer rekurzió. Ha a gyakoriságeloszlásunk  $(a, b, 0)$  típusú és a veszteségeloszlásunk egészértékű, akkor Panjer [1981] bizonyította, hogy az  $S$  teljes veszteség eloszlására teljesül a következő rekurzív képlet:

$$P(S = n) = \sum_{y=1}^n \left( a + \frac{by}{n} \right) P(X = y) P(S = n - y), \text{ ahol } n = 1, 2, \dots, \quad (15)$$

és a rekurziót a  $P(S = 0) = P(X = 0)$  kezdőlépéssel indíthatjuk, ahol  $X$  egy egyedi veszteséget jelöl. Speciálisan,  $\lambda$  paraméterű Poisson gyakoriságeloszlás esetén  $P(S = 0) = e^{-\lambda}$ , továbbá

$$P(S = n) = \frac{\lambda}{n} \sum_{y=1}^n y P(X = y) P(S = n - y), \text{ ahol } n = 1, 2, \dots. \quad (16)$$

Mivel a teljes veszteség eloszlását így egzakt módon számolhatjuk, ezért abból azonnal adódik tetszőleges kvantilis (VaR) értéke. A módszer alkalmazásának veszélye, hogy numerikus hibák léphetnek fel a rekurzió alkalmazása során (részletesebben tárgyalja ezt a kérdéskört Panjer és Willmot (1986)). Továbbá szemben a korábban javasolt folytonos veszteségeloszlásokkal, a rekurzió azt feltételezi, hogy a veszteségeloszlás is diszkrét. Ugyan Panjer et. al. [1992] megmutatta, hogy egy  $f_X$  folytonos egyedi veszteségeloszláshoz tartozó sűrűségfüggvény esetén a teljes eloszlás  $f_S$  sűrűségfüggvénye kielégíti az

$$f_s(x) = P(\eta = 1)f_x(x) + \int_0^x \left(a + \frac{by}{x}\right) f_x(y)f_s(x-y)dy \quad (17)$$

Volterra típusú integrálegyenletet, azonban a gyakorlatban ez nem a legegyszerűbben használható eszköz. Hiszen ennek az eredménynek a használata esetén meg kell oldani az integrálegyenletet, amely számos újabb numerikus kérdést vet fel, komoly szakértelmet igényel (szemben például a fent leírt egyszerű Monte Carlo közelítéssel). Ezért, ha mégis meg akarjuk határozni az teljes veszteségeloszlást, s nem csak a kvantilisre van szükségünk, akkor szerencsésebb a fenti integrálegyenlet helyett a Panjer-féle rekurziós képletet használni még folytonos eloszlások esetén is, azonban ehhez a folytonos veszteségeloszlás diszkretizálása szükséges, hogy ennél fogva alkalmazható legyen a szóban forgó rekurziós képlet Klugman et. al. [2004].

Esetünkben itt is annak a vizsgálata a legfontosabb, hogy a rekurzív módszerrel kapott tőkekövetelmény érték mennyire pontos, illetve mekkora hiba adódhat a módszer használata során. Rekurziós számítások során a kerekítésből, számábrázolásból származó hibák felhalmozódhatnak, és jelentős mértékűvé válhatnak. A rekurzió során minden lépésben valamennyire pontatlan eredményt kapunk, hiszen a végtelen tizedesjegyű számokat is csak végesként tudjuk kezelni. Így azt a kérdést kell megvizsgálni, hogy a számítások során milyen gyorsan növekednek ezek a hibák az ilyen egymásra épülő eljárásoknál. Panjer et. al. [1993] megmutatták, hogy Poisson és negatív binomiális eloszlások esetén a rekurzív formula stabil marad, viszonylag lassan növekedő hibákkal.

#### 7.1.2.1. Diszkretizálás

Amint azt felfezettük a Panjer rekurzió a veszteségeloszlás használatát követeli meg. Nem célunk, hogy a diszkretizálás lépéseit részleteiben leírjuk, ugyanakkor annak érdekében, hogy az ily módon előállított veszteségeloszlás egy érdekes tulajdonságát megismerhessük a diszkretizálás lényegét és módszereit gyorsan összefoglaljuk.

A diszkretizálás során a folytonos veszteségeloszlás értelmezési tartományán diszkrét pontokat határozunk meg, és azokhoz valószínűségeket rendelünk. A diszkretizálás során azt a célt kell kielégítenünk, hogy a diszkrét eloszlás minél jobban reprezentálhassa az eredeti folytonos eloszlást. Természetesen nem mindegy, hogy milyen értelemben kell jól

reprezentálni az eredeti eloszlást. Legyen a diszkrét eloszlás alakja a folytonoshoz hasonló, vagy rendelkezzen ugyanazokkal a momentumokkal, kvantilisekkel? Önmagukban ezek mind lényegtelenek lehetnek, hiszen esetünkben a cél az, hogy az aggregált eloszlás VaR-ja pontos és stabil legyen. A különböző diszkretizálási (kerekítéses és momentum megőrző) eljárások segítségével előállított diszkrét egyedi veszteség eloszlások tehát nem tökéletesen reprodukálják az eredeti folytonos eloszlást, a kérdés csupán, hogy amennyiben egy diszkretizált veszteség eloszlást használunk, mekkora hibát vétünk a tőkekövetelmény megállapítása során.

Panjer et. al. [1983] megmutatták, hogy a folytonos eloszlás megfelelő közelítéséhez az első két momentum illesztése általában elegendő. A diszkretizálásból eredő hiba már alig csökken, ha a modellt további momentumok illesztésével bővítjük. A kerekítéses, első illetve első két momentum illesztésének módszere közötti választásnál néhány szempontot azonban figyelembe kell venni. Így, míg a kerekítéses és az egy momentum illesztésénél kapott valószínűségek mindig pozitívak lesznek, addig a kettő vagy több momentum figyelembe vételénél előfordulhat, hogy negatív valószínűségeket kapunk. Valamint azt is fontos megjegyezni, hogy a kerekítéses és az egy momentum illesztésének módszeréből hasonló nagyságrendű hibák adódhatnak, azonban a második momentum illesztésénél ezek a hibák jelentősen csökkennek. Így ezeket a szempontokat mérlegelni kell, amikor azt vizsgáljuk, hogy melyik a leginkább megfelelő módszer a folytonos eloszlás diszkretizálásához.

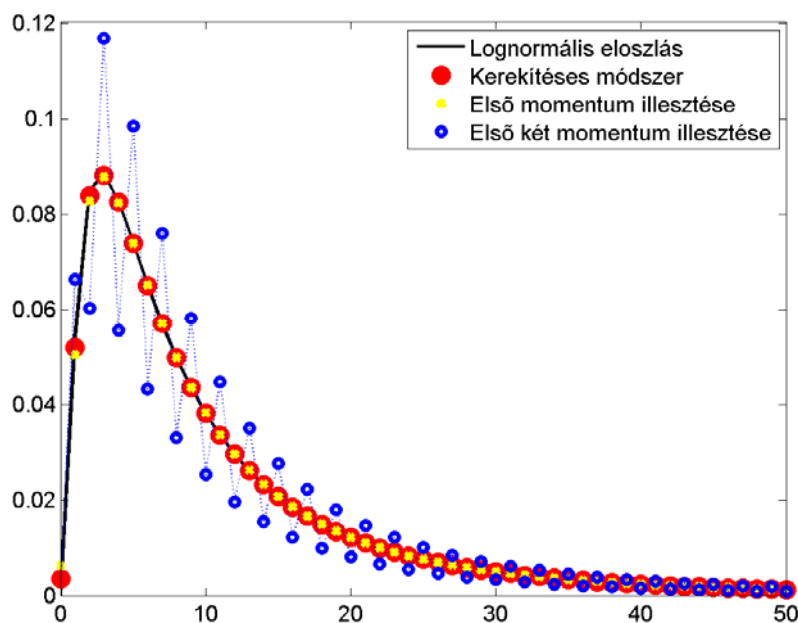
#### *7.1.2.2. Panjer rekurzió és a tőke stabilitása*

Az aggregált veszteségeloszlás VaR-jának viselkedését mind a kerekítéses mind a momentum illesztéses módszerrel teszteltük. Arra kerestük a választ, hogy van-e érdemi különbség az egyes módszerek között. Vizsgálatunk során lognormális, Pareto és exponenciális eloszlásokat használtunk.

A módszerek alkalmazásakor az intervallumok hosszát kezdetben egy egységnek választottuk. Habár tudjuk Panjer [2006] alapján, hogy az egyes diszkretizálási módszerek közötti különbség elsődlegesen akkor jelentős, ha a folytonos eloszlást néhány diszkrét pontból álló eloszlással akarjuk közelíteni, vizsgálataink kezdetén az ebből adódó eltérések minimalizálására törekedtünk. Ennek oka, hogy az ilyen típusú hiba jól csökkenthető a h intervallum kellően kicsi megválasztásával. Minél sűrűbben helyezkednek el az eloszlás

diszkretizált pontjai, annál pontosabb eredményt kapunk, azaz annál jobban tudjuk a diszkretizált valószínűségekkel közelíteni az eredeti folytonos eloszlást. Ugyanakkor megjegyezzük, hogy a számítógépek memóriakapacitása gátat szabhat az egyedi veszteségeloszlás végtelenül finom diszkretizálásának. A kerekítéssel kapott diszkretizált eloszlás valószínűségeit összegezve az elvartaknak megfelelően egyet kaptunk eredményül, és teljesült a kapott értékekre elvart pozitív tulajdonság. Ugyanez igaz, ha csak az első momentum értékét kívánjuk megőrizni a diszkretizálás során.

Az első két momentum illesztésének módszerét használva a diszkretizált valószínűségek ugrálnak, és akár negatívak is lehetnek(!), de a kapott diszkrét eloszlás összességében nagyon jól viselkedett az aggregált káreloszlás meghatározásakor (Nagy et. al. [2009]). A két momentum illesztésével kapott valószínűségekre az első értéktől eltekintve mindig pozitív értékeket kaptunk.



3.ábra: Lognormális eloszlás ( $\mu=2$  és  $\sigma=1$ ) diszkretizálása kerekítéssel és momentum illesztéssel.

Az ábráról jól látszik, hogy a kerekítéssel és az első momentum illesztéssel kapott értékek szinte teljesen egybeesnek. Az első két momentum illesztésének módszerénél pedig az ugráló valószínűségeket lehet jól megfigyelni.

Megnéztük azt is, hogy a rekurzióval előállított aggregált eloszlás (Poisson  $\lambda=10$  – lognormális  $\mu=2$ ,  $\sigma=1$ ) várható értéke és szórása mennyiben különbözik az elméleti, valamint a Monte-Carlo szimulációval kapott értékektől. A Monte-Carlo szimulációs esetben 100 000-

es szimulációs szám mellett 1000-szer határoztuk az aggregált eloszlást, és azokból átlagoltuk ki az eloszlás jellemzőit. A kapott eredményeket az alábbi táblázat foglalja össze.

	Elméleti	Monte-Carlo	Kerekítéses módszer	Első momentum illesztése	Első két momentum illesztése
Várható érték	121,8249	121,8306	121,8306	121,8262	121,8249
Szórás	63,5160	63,5226	63,5218	63,5283	63,5156

3. táblázat: A különböző módszerekkel meghatározott aggregált káreloszlások első két momentumai.

A táblázatból látható, hogy a különböző módszerekkel végzett diszkretizálások során az eloszlások mind várható értéküket, mind szórásukat rendkívül pontosan megőrizték. Már a kerekítéses módszer is nagyon jól közelítette az elméleti várható értéket és szórást (itt újra megjegyezzük, hogy a diszkrét pontok kellően sűrűn helyezkedtek el). Az első két momentum megőrzésével kapott eredmények pedig már szinte minden tizedesre megegyeznek a várt értékekkel. Ezek után megváltoztattuk a lognormális és Poisson eloszlások paramétereit ( $\mu = [0,5; 1; 2; 5]; \sigma = [1; 2]; \lambda = [0,05; 1; 10; 100]$ ). Az így kapott eredményekre ugyanúgy teljesültek a fenti megállapítások, azaz a különböző diszkretizálási módszerek segítségével meghatározott aggregált eloszlások meglehetősen pontosan visszaadják az elméleti eloszlás megfelelő momentumait. Ezek közül is az első két momentum megőrzésével kapott eredmények voltak legközelebb az elméleti értékekhez, de különbség minden esetben elhanyagolható volt.

Az aggregált eloszlásokból meghatároztuk a kockázatos értéket is különböző konfidenciaszintek mellett, ezeket a következő táblázat foglalja össze.

	Monte-Carlo	Kerekítési módszer	Első momentum illesztése	Első két momentum illesztése
VaR 90	203.2	204	204	204
VaR 95	238.5	240	240	240
VaR 99	322.8	324	324	324
VaR 99.5	362.2	363	363	363
VaR 99.9	467.5	468	468	468

4. táblázat: A különböző diszkretizálási módszerekkel meghatározott aggregált káreloszlások VaR értékei. Viszonyításképpen a Monte Carlo szimulációval kapott értékeket is feltüntetjük.

A táblázatból látható, hogy a három (kerekítési, első momentum illesztése, első két momentum illesztése) módszer egész számra kerekítve azonos eredményre vezetett, a kockázatos értékek egymással teljesen megegyeznek. A korábban leírt módon módosított  $\mu$ ,  $\sigma$  és  $\lambda$  paraméterekkel meghatározott aggregált eloszlásoknál is teljesült a fenti állítás. Azaz a különböző diszkretizálási módszerekkel kiszámított VaR értékek nem különböztek lényegesen egymástól.

A fenti tesztet (diszkretizálás közelítéssel és lokális momentum illesztéssel, majd a rekurzió alkalmazása és VaR számítás) megismételtük Pareto és exponenciális egyedi veszteségeloszlásokkal is, és a korábban vázoltakkal megegyező eredményre jutottunk. Összességében, a kockázatos érték szempontjából a három diszkretizálási módszer ugyanarra az eredményre vezetett, ugyanakkor nem szabad arról megfeledkeznünk, hogy ez egy olyan modellezési környezetben történt, ahol a folytonos eloszlás diszkretizálása egy egységnyi intervallumonként történt. Természetesen a diszkretizálási intervallum növelésével a momentum illesztési módszerek egyre jobb eredményt adnak.

A Panjer rekurzió gyors lehetőséget biztosít az aggregált eloszlás és annak jellemzőinek a meghatározására, ugyanakkor meg sem közelíti a szimulációs módszerek rugalmasságát. Így nem nyilvánvaló hogyan lehetne pl. bevonni a biztosítások hatását a modellbe, illetve figyelembe venni az egyes veszteségek, vagy kockázati osztályok közötti korrelációt.

### 7.1.3. Fourier transzformáció



Az összetett eloszlás meghatározására egy további, szintén nem triviális mód vezet az eloszlás momentumgeneráló és karakterisztikus függvényének meghatározásán keresztül. Ehhez tudnunk kell, hogy a teljes eloszlás momentumgeneráló függvénye meghatározható a gyakoriság generátorfüggvénye és az egyedi veszteségek momentumgeneráló függvénye segítségével (amennyiben léteznek):

$$G_S(y) = g_\eta(G_X(y)), \quad y \in R,$$

ahol  $G_X$  és  $G_S$  az egyedi illetve a teljes veszteség eloszlásainak momentumgeneráló függvényét jelöli, míg  $g_\eta$  a gyakoriságeloszlás generátorfüggvénye. Ugyanilyen összefüggés adódik karakterisztikus függvények esetére is. Utóbbi előnye, hogy mindig létezik, továbbá mindig egyértelmű. Ha a fenti módon meghatároztuk a teljes veszteség karakterisztikus függvényét, akkor a gyors Fourier transzformáció (Fast Fourier Transform, FFT) módszerét használhatjuk annak érdekében, hogy a karakterisztikus függvényből megkapjuk (becsüljük) a szóban forgó eloszlást (sűrűségfüggvényt).

Az aggregált veszteségeloszlás meghatározásának lépései:

1. A veszteség eloszlás diszkrétizálása.
2. A gyors Fourier-transzformáció alkalmazása a diszkrét veszteségeloszlásra.
3. Az összetett eloszlás karakterisztikus függvényének meghatározása. A fenti képlet alapján alkalmazzuk a gyakoriságeloszlás generátorfüggvényét a már diszkrét veszteségeloszlás karakterisztikus függvényére. Ez adja meg az aggregált eloszlás karakterisztikus függvényét, ami az összetett eloszlás diszkrét Fourier-transzformációjának felel meg.

4. Az aggregált eloszlás karakterisztikus függvényre az inverz gyors Fourier transzformációt alkalmazva megkapjuk a keresett összetett eloszlást diszkrét alakban.
5. Az aggregált eloszlásból határozzuk meg a különböző (de elsődlegesen a  $p=0,999$ ) konfidenciaszintekhez tartozó kockázatos értékeket, és az eloszlás egyéb jellemzőit.

A Fourier transzformációs eljárás kritikus eleme az egyedi veszteségeloszlás diszkretizálása lehet. Ugyanakkor itt is igaz, hogy amennyiben a diszkretizálást kellően finom lépésközönként végezzük, akkor még a kerekítéses módszer is megállja a helyét, és az eljárás pontos eredményt ad. Természetesen ezt fokozhatjuk a momentum illesztéses módszerekkel. A nagyon finom lépésköz használatának ebben az esetben is a memória szabhat gátat, a köztes pontok száma legyen kisebb mint  $2^{24}$ . Ebben az esetben ez egy kiváló eszköz az aggregált eloszlás és annak jellemzőinek a meghatározására. További előnye, hogy gyors, ugyanakkor itt sem egyértelmű, hogyan lehetne bevonni a biztosítások hatását a modellbe. Mivel egydimenziós eljárásról van szó, ezért a dimenziók ex-post aggregációja is szükséges (Mignola et. al. [2006]).

## 7.2. Single Loss Approximation és tőkeszámítás

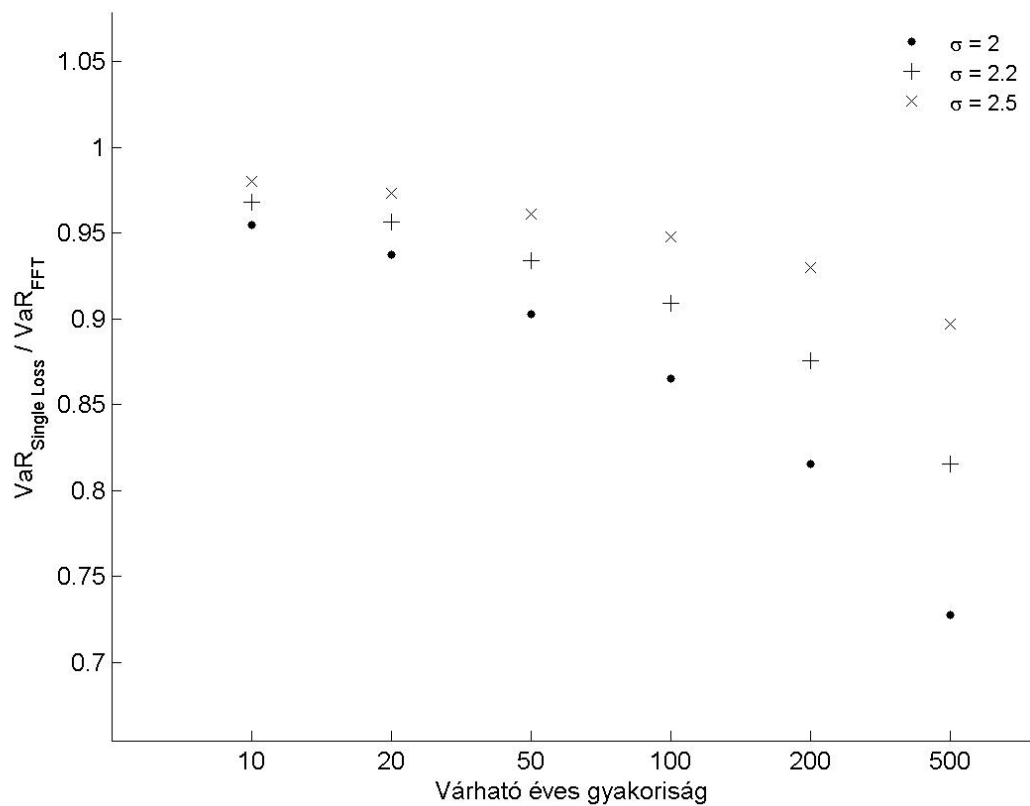
A könnyebb áttekinthetőség kedvéért itt újra vázoljuk, hogy szubexponenciális eloszlások esetén (pl. lognormális, Weibull, Pareto, log-gamma eloszlások) a tőkekövetelmény az alábbi módon is meghatározható:

$$VaR(p) = F_s^{-1}(p) \approx F^{-1}\left(1 - \frac{1-p}{E(N)}\right). \quad (18)$$

Amint azt már említettük a fenti képlet direkt kapcsolatot teremt az az aggregált eloszlás ( $F_s$ )  $p$  konfidenciaszinthez tartozó kvantilise és az  $F$  egyedi veszteségeloszlás  $1 - \frac{1-p}{E(N)}$  konfidenciaszinthez tartozó kvantilise között.

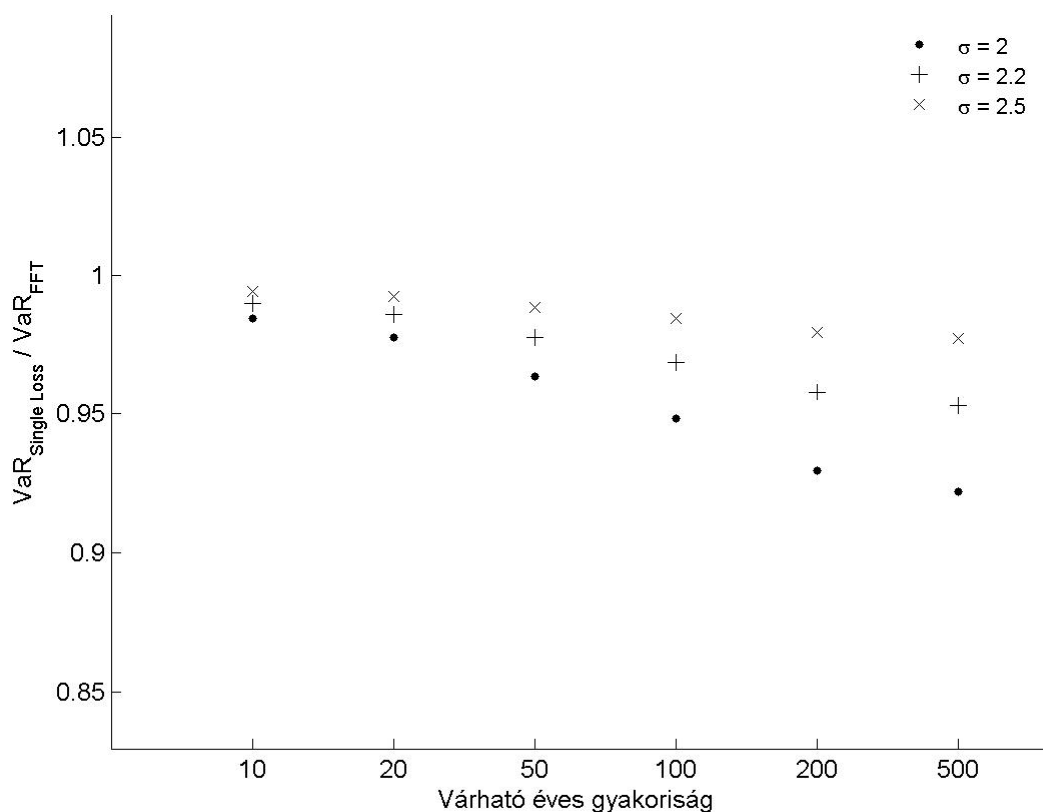
Azt, hogy a Single Loss Approximation, mint tőkebecslési módszer mennyire ad pontos becslést, avagy az eredmények mennyire állnak közel egy LDA-val számított tőkéhez az

alábbi ábrával szemléltetjük. Mindkét módszer (LDA, SLA) esetében az igazi, becslési hibától mentes paraméterekkel számoltuk a tőkét.



17. ábra: A Single Loss Approximation módszerrel, valamint Fourier transzformációval meghatározott aggregált eloszlás 0,999-es valószínűségéhez tartozó VaR értékek hányadosa a várható éves gyakoriság függvényében. Az éves gyakoriság eloszlás Poisson eloszlású az x tengelyen feltüntetett várható értékkel, míg az alkalmazott veszteség eloszlás lognormális  $\mu = 10$  rögzített paraméterrel és  $\sigma = 2$  ;  $\sigma = 2,2$  ;  $\sigma = 2,5$  értékekkel.

Jól látható a 17. ábra alapján, hogy az éves gyakoriság növekedésével az SLA módszer egyre kisebb hatékonysággal becsli a tőkét. A másik amit leolvashatunk a fenti ábráról, hogy amennyiben a veszteség eloszlás kevésbé vastag szélű (esetünkben ahol  $\sigma$  kicsi), ott az SLA megbízhatósága szintén gyenge. Ennek oka a modell feltételekben keresendő, hiszen Böcker et. al. [2005] megmutatta, hogy az összefüggés csak limeszben teljesül ( $p \rightarrow 1$ ), azaz 1-hez igen közeli valószínűség mellett kell a kvantilist meghatározni. A tőkebecsléshez használandó valószínűséget viszont a szabályozó írja elő, így a modellalkotó erről már nem rendelkezhet. A 18. ábra megmutatja, hogy  $p=0,9999$ -es valószínűség mellett hogyan alakul a tőkebecslés a valós tőkéhez képest. Amint az vártuk, egy nagyságrenddel közelebb kerülve a  $p=1$ -hez javul az SLA becslés hatékonysága, ami az 17. ábra és a 18. ábra összevetéséből derül ki. A két ábra közötti különbség mindössze a VaR számításához alkalmazott valószínűség.



18. ábra: A Single Loss Approximation módszerrel, valamint Fast Fourier transzformációval meghatározott aggregált eloszlás 0,9999-es valószínűségéhez tartozó VaR értékek hányadosa a várható éves gyakoriság függvényében.<sup>25</sup>

Amint arra már korábban is utaltunk az intézménynek nincs ráhatása az alkalmazott valószínűségekre a tőkeképzés során, így az SLA-ból adódó pontosabb tőkeképzés érdekében nem használhat  $p=0,999$ -től eltérő konfidencia szintet (legalábbis 1-es Pillérben). Az empirikus vizsgálatok rámutattak (lásd pl. Aue et. al. [2006]), hogy a gyakorlatban az SLA eredményeitől eltérően a tőke szempontjából nem csak a legnagyobb veszteség számít, hanem a legnagyobb veszteségek számítanak, azaz a top néhány veszteség határozza meg a tőkét. Ennek oka, hogy a gyakorlatban alkalmazott eloszlások heavy tail jellege sem feltétlenül erős (vagy extrém erős), így nincs az az egyetlen hatalmas veszteség, ami egyértelműen dominálná a tőkét. Ez mindenképp megfontolást igényel a Single Loss Approximation használata

<sup>25</sup> Az éves gyakoriság eloszlás Poisson eloszlású az x tengelyen feltüntetett várható értékkel, míg az alkalmazott veszteség eloszlás lognormális  $\mu = 10$  rögzített paraméterrel és  $\sigma = 2$ ;  $\sigma = 2,2$ ;  $\sigma = 2,5$  értékekkel.

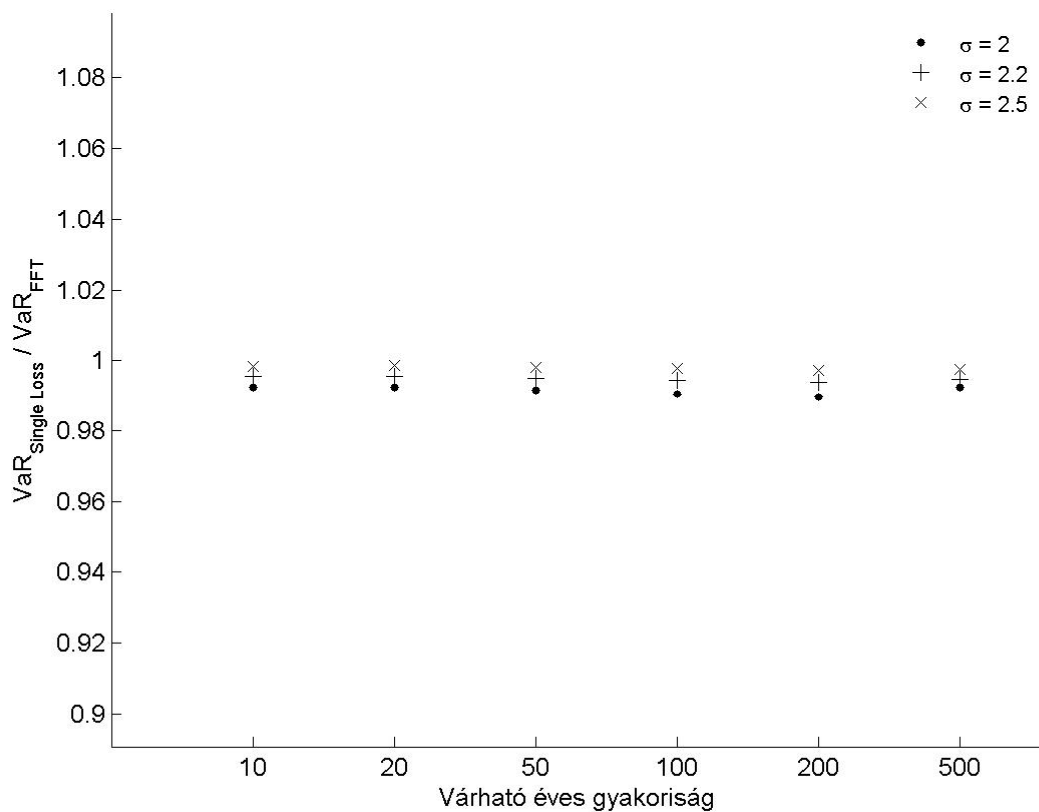
szempontjából, ugyanakkor Böcker et. al. [2006] a Single Loss Approximation egy finomított változatát publikálta. Eszerint

$$VaR(p) = F^{\leftarrow} \left( 1 - \frac{1-p}{EN} (1 + o(1)) \right) + E(N-1)E(X), \quad p \rightarrow 1. \quad (19)$$

Ez alapján a Single Loss Approximation várható értékkel korrigált változata szerinti kvantilis érték:

$$VaR(p) = F_s^{-1}(p) \approx F^{-1} \left( 1 - \frac{1-p}{E(N)} \right) + E(N-1)E(X), \quad (20)$$

ahol  $E(X)$  a veszteség eloszlás várható értékét jelenti.



19. ábra: A várható értékkel korrigált Single Loss Approximation, valamint Fast Fourier transzformációval meghatározott aggregált eloszlás 0,9999-es valószínűségéhez tartozó VaR értékek hányadosa a várható éves gyakoriság függvényében.<sup>26</sup>

A 19. ábra jól mutatja, hogy a várható értékkel korrigált Single Loss Approximation lényegesen pontosabb becslést ad a tőkére, mint a korrekció nélküli változata. Ugyanakkor ez egy „heurisztikus” választáson alapuló korrekciója Böcker et. al. [2005] eredményének, Degen [2010] azonban egy még precízebb korrekcióját adja a Single Loss Approximationnek. Eszerint a veszteségeloszlás heavy tail jellegétől függően három esetet különböztet meg:

1. heavy tail eloszlás, véges várható értékkel;
2. heavy tail eloszlás, nem véges várható értékkel;
3. semi heavy tail eloszlások (pl. lognormális).

Az (1)-es esetben Degen javaslata a következő:

$$VaR(p) = F^{\leftarrow} \left( 1 - \frac{1-p}{EN} \right) + E(N-1+d)E(X), \quad (21)$$

ahol  $d = \frac{D^2(N)}{E(N)}$ .

Poisson gyakoriság eloszlás eloszlás esetén  $d=1$ . Felhívjuk a figyelmet, arra, hogy pl. binomiális gyakoriság eloszlás esetén  $d < 1$ , tehát a korrekciós tag negatív is lehet.

Degen [2010] ugyan precíz közelítést ad olyan veszteségeloszlásokra / esetekre is (2), amikor az eloszlás várható értéke nem létezik, ugyanakkor ezek közel sem tartoznak a main stream eloszlások közé. A tőke modellezése során fontos a matematikai modell közgazdasági interpretációja, márpedig egy olyan állítás, hogy a várható veszteség végtelen, miközben a pénzügyi intézmény minden folyamata legfeljebb véges mértékű veszteséget generálhat

---

<sup>26</sup> Az éves gyakoriság eloszlás Poisson eloszlású az x tengelyen feltüntetett várható értékkel, míg az alkalmazott veszteség eloszlás lognormális  $\mu = 10$  rögzített paraméterrel és  $\sigma = 2$ ;  $\sigma = 2,2$ ;  $\sigma = 2,5$  értékekkel.

nehezen elfogadható. Az erre az esetre adott korrekciós javaslat megtekinthető Degen [2010] publikációjában.

Semi heavy tail eloszlások (3) esetén Degen [2010] nem tud javaslatot adni, így továbbra is (Böcker et. al. [2006]) javaslatát tartjuk irányadónak.

### 7.2.1. *Single Loss Approximation és az extrapolációs probléma* (kutatási hipotézis)

Az SLA megközelítés segít számunkra megérteni, hogy miért is annyira problémás a működési kockázati tőke becslése, miért jellemzi azt nagy instabilitás. Az extrapolációs probléma, azaz hogy függetlenül az éves gyakoriság várható értékétől mindenkinek 200-szor több adatra lenne szüksége, felveti azt a kérdést, hogy pontosabban tudják-e azok a pénzüzetek becsülni a tőkéjüket, melyek több veszteség adattal rendelkeznek? Akinek csupán évente 10 adata van, annak az egyedi veszteségeloszlás 0,9999-es kvantilisét jelenti a tőkét, viszont aki átlagosan 1000 veszteségadatot gyűjt évente, annak a 0,999999-es valószínűséghez tartozó kvantilisét kell meghatároznia. Két ellentétes hatás áll szemben egymással, melyek meghatározzák a tőke becslési hibáját. Egyrészt a nagyobb minta kedvező a becslés szempontjából, ugyanakkor az egyre távolabbi (magasabb valószínűséghez tartozó) kvantilisek becslése növekvő bizonytalanságot hordoz.

Fentiek alapján a két egymással ellentétes hatás következtében nem nyilvánvaló, hogy a sok adattal rendelkező hitelintézetek valóban kisebb, vagy lényegesen kisebb hibával becslik-e tőkéjüket.

Hogy közelebb kerüljünk ezen felvetés megválaszolásához, bizonyos feltételek között meghatározzuk a tőke eloszlását és annak főbb jellemzőit, köztük sztenderd hibáját. Mindez egy praktikus modellezési környezetet teremt számunkra, melyből számos következtetést vonhatunk le a működési kockázatok természetéről.

### 7.2.2. *A tőkebecslés hibája lognormális veszteségeloszlás esetén* (önálló eredmény)

A lognormális eloszlás és a normális eloszlás között függvényszerű kapcsolat van azaz  $\eta = e^{\xi}$ , ahol  $\eta$  lognormális  $\xi$  pedig normális eloszlású valószínűségi változó. Tetszőleges

normális eloszlású változó tetszőleges  $q$  kvantilise előállítható az eloszlás várható értékének ( $\mu$ ) és szórásának ( $\sigma$ ) a függvényében,  $q = \mu + z\sigma$ , ahol  $z$  a sztenderd normális eloszlás ugyanazon valószínűséghez tartozó kvantilise. A fentiek tekintetében és a normális eloszlás becsült paramétereinek ( $\hat{\mu}, \hat{\sigma}^2$ ) eloszlásából limeszben analitikusan meghatározható, míg véges minta esetén közelíthető a lognormális eloszlású változó kvantilisének eloszlása.

Az  $n$  elemű mintából Maximum Likelihood módszerrel becsült normális eloszlás

paramétereinek várható értéke:  $\begin{bmatrix} \mu \\ \sigma^2 \end{bmatrix}$ ,

variancia-kovariancia mátrixa:  $\begin{bmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{2\sigma^4}{n} \end{bmatrix}$ .

Lásd pl. Lynch [2007].

Maximum Likelihood becslés esetén a becsült paraméterek korrelálatlanok és aszimptotikusan

normális eloszlásúak, így  $\hat{\mu} \approx N(\mu, \frac{\sigma^2}{n})$  és  $\hat{\sigma}^2 \approx N(\sigma^2, \frac{2\sigma^4}{n})$ .

Véges kis ( $n$ ) minta esetén  $\hat{\sigma}^2, \hat{\sigma}$  eloszlása nem normális eloszlású (Ahn et. al. [2003]) ugyanakkor a paraméterek korrelálatlansága továbbra is fennáll.  $\hat{\sigma}$  becslésének standard

hibája  $\frac{\sigma}{\sqrt{2(n-1)}} \left(1 + O\left(\frac{1}{n}\right)\right)$ . A következő levezetésben feltételezzük, hogy  $\hat{\sigma}$  normális

eloszlású  $\sigma$  várható értékkel és  $\frac{\sigma}{\sqrt{2n}}$  szórással (Ahn et. al. [2003] már  $n > 10$  esetén 1-nek

veszik  $\left(1 + O\left(\frac{1}{n}\right)\right)$  értékét.)

A lognormális eloszlás kvantilisére teljesül, hogy  $q_L = e^{q_N} = e^{\mu + \alpha\sigma}$ , ahol  $\mu$  és  $\sigma$  a normális eloszlású változó paramétere, és egyben a várható érték és szórás jellemzői.  $\mu$  és  $\sigma$  becslése

során tudjuk, hogy  $\hat{\mu}$  várható értéke:  $\mu - \frac{\sigma^2}{n}$  varianciával, míg  $\hat{\sigma}$  várható értéke:  $\sigma$  és

közelítőleg  $\frac{\sigma^2}{2n}$  varianciával, miközben  $\hat{\mu}$  és  $\hat{\sigma}$  eloszlásáról feltételezzük a normalitást és a

korrelálatlanságot.



Mivel  $\hat{\mu}$  és  $\hat{\sigma}$  normális eloszlású változók és korrelálatlanok (feltételezésünk szerint), ezért azok lineáris kombinációja  $\hat{\mu} + \alpha\hat{\sigma}$  is normális eloszlású lesz  $q_N$  várható értékkel és  $\alpha^2 \frac{\sigma^2}{2n} + \frac{\sigma^2}{n} = \frac{\sigma^2(\alpha^2 + 2)}{2n}$  varianciával.

A lognormális eloszlás kvantilise tehát  $q_L = e^{q_N} = e^{\mu + \alpha\sigma}$  lognormális eloszlást követ, hiszen az exponenciális függvény kitevőjébe egy normális eloszlású változót tettünk. Így már könnyen meghatározhatjuk azt is, hogy ennek a  $\hat{q}_L$  lognormális eloszlású változónak milyen paraméterű az eloszlása, mekkora a sztenderd hibája a becslés során?

A lognormális eloszlás (MU, SIGMA) várható értéke  $E = e^{\frac{MU + SIGMA^2}{2}}$ , míg varianciája  $D^2 = (e^{SIGMA^2} - 1)e^{2MU + SIGMA^2}$ . Esetünkben,  $\hat{q}_L$  olyan lognormális eloszlású, ahol  $MU = \mu + \alpha\sigma$  és  $SIGMA^2 = \frac{\sigma^2(\alpha^2 + 2)}{2n}$ .

Behelyettesítve kapjuk, hogy

$$E(\hat{q}_L) = e^{\frac{MU + SIGMA^2}{2}} = e^{\mu + \alpha\sigma + \frac{\sigma^2(\alpha^2 + 2)}{2n}} \quad (22)$$

$$D^2(\hat{q}_L) = (e^{SIGMA^2} - 1)e^{2MU + SIGMA^2} = \left( e^{\frac{\sigma^2(\alpha^2 + 2)}{2n}} - 1 \right) \left( e^{2(\mu + \alpha\sigma) + \frac{\sigma^2(\alpha^2 + 2)}{2n}} \right). \quad (23)$$

Gyökvonás után adódik:

$$\sigma_0 = \left( (e^{SIGMA^2} - 1)e^{2MU + SIGMA^2} \right)^{\frac{1}{2}} = \left( \left( e^{\frac{\sigma^2(\alpha^2 + 2)}{2n}} - 1 \right) \left( e^{2(\mu + \alpha\sigma) + \frac{\sigma^2(\alpha^2 + 2)}{2n}} \right) \right)^{\frac{1}{2}}. \quad (24)$$

A fenti képlet tehát megadja a lognormális eloszlású változó tetszőleges kvantilisének sztenderd hibáját ( $\sigma_0$ ). Természetesen a sztenderd hibán túl a teljes eloszlást ismerjük, így annak számos tulajdonságát megismerhetjük. Felhívjuk a figyelmet arra, hogy az így számított hiba, csak az egyedi veszteségeloszlás becsléséből származó hibát foglalja magában, hiszen a gyakoriság várható értékének becsléséből származó hibával foglalkoztunk. Annak hatása valójában  $\alpha$ -n keresztül érvényesülne, hiszen az SLA keretein belül az eredeti

szabályozói valószínűséget ( $p=0,999$ ) a gyakoriság várható értékével kell skálázni ( $1 - \frac{1-p}{E(n)}$ ), ami épp az  $\alpha$  stabilitását befolyásolja.

A levezetés eredményét kombinálva a Single Loss Approximation-nel (SLA), miszerint  $VaR = F_s^{-1}(p) \approx F^{-1}(1 - \frac{1-p}{E(n)})$  választ kaphatunk arra, hogy különböző várható gyakoriságú működési kockázati események tőkeszükségletét (feltéve, hogy maga a veszteség lognormális eloszlású) mekkora hibával tudjuk megadni, azaz a mintaelemszám növekedése hogyan hat a számított tőke stabilitására. Valóban fennáll-e az a hipotézis, hogy az extrapolációs probléma miatt (minden gyakoriság mellett 200-szoros extrapoláció szükséges) kis és nagy bankok, azaz sok és kevés adattal rendelkező intézmények ugyanakkora hibával becslik a tőkésükét?

Mindenekelőtt azonban szükségesnek tartjuk a fenti – tőke eloszlására vonatkozó eredmények visszamérését, hiszen több ponton éltünk közelítéssel a levezetés során. Ennek megfelelően a következő visszaméréseket végeztük el:

1. Adott éves gyakorisághoz és veszteségeloszláshoz tartozó aggregált eloszlás VaR-jának SLA-val történő számítása R-szer, oly módon, hogy minden (1,2...R) alkalommal csak az egyedi veszteségeloszlás paramétereit becsüljük ( $5\lambda$  nagyságú mintából). Az így kapott R darab tőkeértéket –  $VaR_1$  felhasználva sztenderd hibát –  $\sigma_1$  számolunk, amit összevetünk az analitikusan számított  $VaR_0$  –hoz tartozó  $\sigma_0$  –val.
2. Adott éves gyakorisághoz és veszteségeloszláshoz tartozó aggregált eloszlás VaR-jának SLA-val történő számítása R-szer, oly módon, hogy minden (1,2...R) alkalommal a gyakoriság eloszlás és az egyedi veszteségeloszlás paramétereit egyaránt becsüljük (5 éves idősről). Az így kapott R darab tőkeértéket –  $VaR_2$  felhasználva sztenderd hibát –  $\sigma_2$  számolunk, amit összevetünk az analitikusan számított  $\sigma_0$  –al.

A következő táblázatok a levezetésből kapott főbb jellemzőket vetik össze az (1) – ben megfogalmazott szimulációs eredményekkel.

Éves gyako-	VaR (SLA alapján igazi	VaR várható értéke	VaR várható értéke	VaR szórása szimulációk-	VaR szórása levezetés	Várható értékek	Szórások aránya -
-------------	------------------------	--------------------	--------------------	--------------------------	-----------------------	-----------------	-------------------

riság	paraméterek- kel)	szimulációk alapján	levezetés alapján	ból számítva	alapján	aránya - Lev/szim	Lev/szim
10	37.431.867	46.352.647	51.374.610	45.312.087	48.295.145	1,11	1,07
20	52.755.773	58.969.464	62.617.217	38.501.637	40.035.576	1,06	1,04
50	81.405.981	84.769.387	87.788.833	34.097.758	35.439.196	1,04	1,04
100	111.527.481	114.774.311	116.122.959	33.736.684	33.677.183	1,01	1,00
500	223.018.375	224.623.713	225.103.457	31.007.660	30.853.267	1,00	1,00
1000	296.289.809	297.444.092	297.750.852	29.709.149	29.605.733	1,00	1,00

13. táblázat: Adott éves gyakoriságú (1. oszlop)  $\mu=10$  és  $\sigma=2$  paraméterű veszteségeloszláshoz tartozó aggregált eloszlás VaR-jának számítása az igazi paraméterekkel, valamint az (1)-es pontban megfogalmazott feltételek mellett számolt VaR értékek várható értéke és szórása, valamint ezek páronkénti aránya.

Éves gyako- riság	VaR (SLA alapján igazi paraméterek- kel)	VaR várható értéke szimulációk alapján	VaR várható értéke levezetés alapján	VaR szórása szimulációk- ból számítva	VaR szórása levezetés alapján	Várható értékek aránya - Lev/szim	Szórások aránya - Lev/szim
10	78.753.735	100.992.319	115.519.337	108.475.102	123.968.306	1,14	1,14
20	114.868.913	132.950.607	141.336.804	100.017.153	101.322.719	1,06	1,01
50	185.108.971	198.332.188	202.812.561	88.556.072	90.796.686	1,02	1,03
100	261.713.185	272.177.459	274.817.510	88.664.644	88.048.797	1,01	0,99
500	560.893.271	566.328.717	567.244.725	85.443.070	85.606.739	1,00	1,00
1000	766.644.287	770.166.595	771.220.960	84.983.905	84.395.451	1,00	0,99

14. táblázat: Adott éves gyakoriságú (1. oszlop)  $\mu=10$  és  $\sigma=2,2$  paraméterű veszteségeloszláshoz tartozó aggregált eloszlás VaR-jának számítása az igazi paraméterekkel, valamint az (1)-es pontban megfogalmazott feltételek mellett számolt VaR értékek várható értéke és szórása, valamint ezek páronkénti aránya.

Éves gyako- riság	VaR (SLA alapján igazi paraméterek- kel)	VaR várható értéke szimulációk alapján	VaR várható értéke levezetés alapján	VaR szórása szimulációk- ból számítva	VaR szórása levezetés alapján	Várható értékek aránya - Lev/szim	Szórások aránya - Lev/szim
10	240.334.248	341.907.655	394.158.069	453.983.430	512.365.335	1,15	1,13
20	369.063.663	449.183.767	482.379.031	384.240.057	405.984.312	1,07	1,06
50	634.722.934	693.941.769	714.179.806	361.329.294	368.365.907	1,03	1,02
100	940.786.687	979.466.284	1.002.054.798	362.861.851	367.482.140	1,02	1,01
500	2.237.121.665	2.262.891.174	2.269.888.244	389.625.155	389.919.247	1,00	1,00

1000	3.190.874.238	3.206.811.818	3.215.493.631	397.362.083	400.205.058	1,00	1,01
------	---------------	---------------	---------------	-------------	-------------	------	------

15. táblázat: Adott éves gyakoriságú (1. oszlop)  $\mu=10$  és  $\sigma=2,5$  paraméterű veszteségeloszláshoz tartozó aggregált eloszlás VaR-jának számítása az igazi paraméterekkel, valamint az (1)-es pontban megfogalmazott feltételek mellett számolt VaR értékek várható értéke és szórása, valamint ezek páronkénti aránya.

Éves gyakoriság	VaR (SLA alapján igazi paraméterekkel)	VaR várható értéke szimulációk alapján	VaR várható értéke levezetés alapján	VaR szórása szimulációkból számítva	VaR szórása levezetés alapján	Várható értékek aránya - Lev/szi m	Szórás arány Lev/szi m
10	1.543.084.938	2.772.474.835	3.146.159.121	6.240.467.341	5.590.093.604	1,13	0,90
20	2.581.859.362	3.549.115.995	3.796.523.160	3.981.177.984	4.092.953.362	1,07	1,03
50	4.948.938.587	5.681.414.266	5.865.077.106	3.681.239.919	3.730.231.033	1,03	1,01
100	7.935.977.580	8.515.033.043	8.690.742.386	3.852.256.921	3.879.408.885	1,02	1,01
500	22.440.811.654	22.804.582.635	22.915.639.213	4.700.977.257	4.738.936.680	1,00	1,01
1000	34.363.917.002	34.634.342.625	34.746.361.403	5.231.156.964	5.198.309.654	1,00	0,99

16. táblázat: Adott éves gyakoriságú (1. oszlop)  $\mu=10$  és  $\sigma=3$  paraméterű veszteségeloszláshoz tartozó aggregált eloszlás VaR-jának számítása az igazi paraméterekkel, valamint az (1)-es pontban megfogalmazott feltételek mellett számolt VaR értékek várható értéke és szórása, valamint ezek páronkénti aránya.

A (2) – es pont feltételei melletti visszaméréseket nem tesszük teljes egészében közzé, hiszen annak eredményei alig különböznek az (1) – es pontbeli eredményektől.

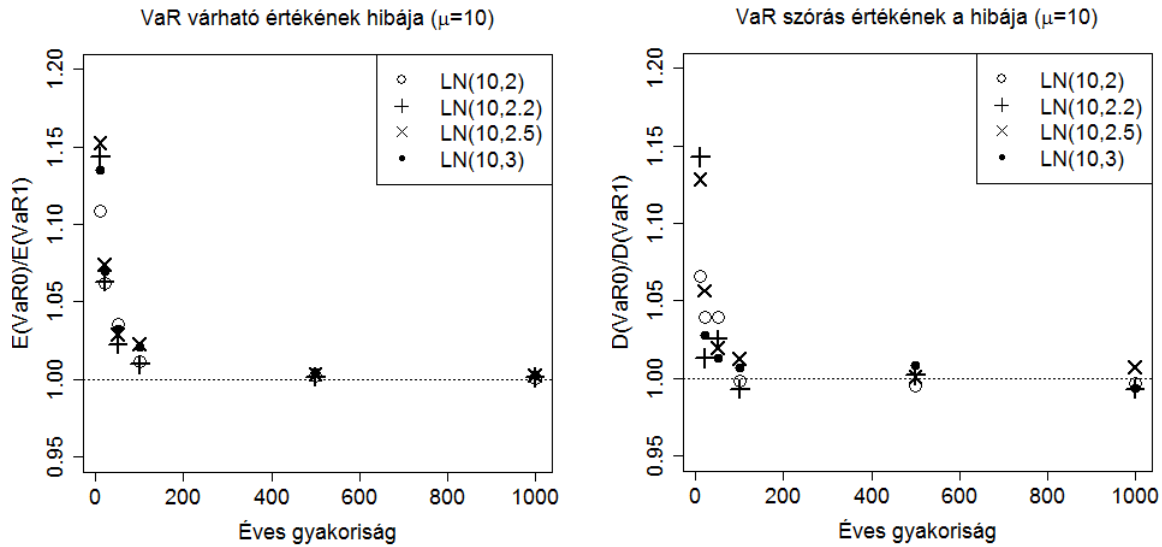
Éves gyakoriság	VaR (SLA alapján igazi paraméterekkel)	VaR várható értéke szimulációk alapján	VaR várható értéke levezetés alapján	VaR szórása szimulációkból számítva	VaR szórása levezetés alapján	Várható értékek aránya - Lev/szi m	Szórás arány Lev/szi m
10	37.431.867	45.339.507	51.374.610	44.591.900	48.295.145	1,13	1,08
20	52.755.773	59.188.803	62.617.217	37.932.624	40.035.576	1,06	1,06
50	81.405.981	86.042.449	87.788.833	34.994.247	35.439.196	1,02	1,01
100	111.527.481	114.792.075	116.122.959	33.129.299	33.677.183	1,01	1,02
500	223.018.375	224.424.785	225.103.457	30.861.128	30.853.267	1,00	1,00
1000	296.289.809	297.279.257	297.750.852	29.500.740	29.605.733	1,00	1,00

17. táblázat: Adott éves gyakoriságú (1. oszlop)  $\mu=10$  és  $\sigma=2$  paraméterű veszteségeloszláshoz tartozó aggregált eloszlás VaR-jának számítása az igazi paraméterekkel, valamint az (2)-es pontban megfogalmazott feltételek mellett számolt VaR értékek várható értéke és szórása, valamint ezek páronkénti aránya.

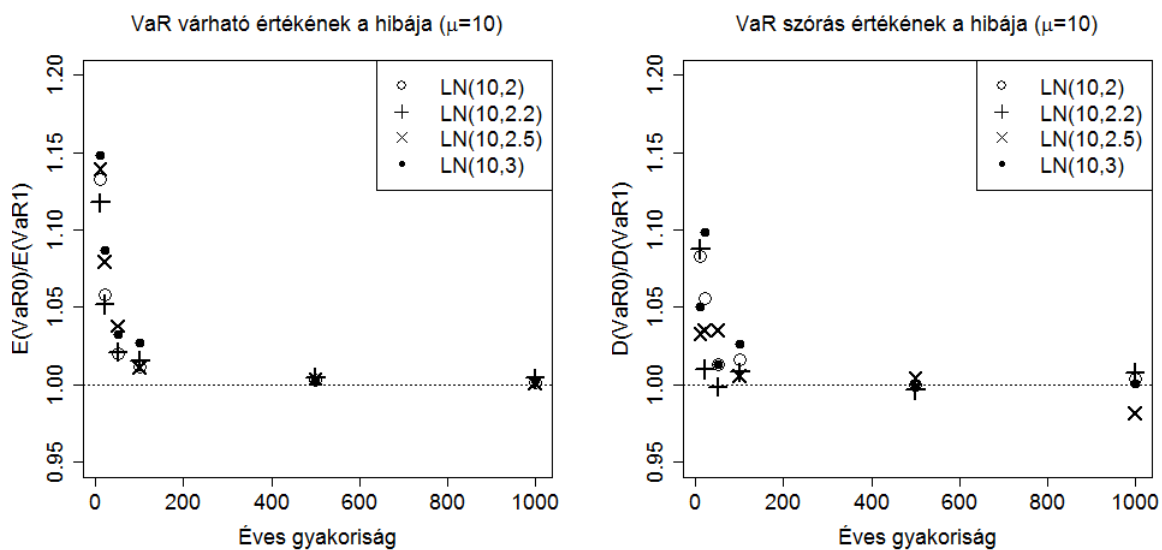
Éves gyakoriság	VaR (SLA alapján igazi paraméterekkel)	VaR várható értéke szimulációk alapján	VaR várható értéke levezetés alapján	VaR szórása szimulációkból számítva	VaR szórása levezetés alapján	Várható értékek aránya - Lev/szim	Szórás arány Lev/szim
10	1 543 084 938	2 740 773 199	3 146 159 121	5 322 079 445	5 590 093 604	1,15	1,05
20	2 581 859 362	3 491 992 449	3 796 523 160	3 726 353 169	4 092 953 362	1,09	1,10
50	4 948 938 587	5 682 266 574	5 865 077 106	3 682 856 268	3 730 231 033	1,03	1,01
100	7 935 977 580	8 464 343 972	8 690 742 386	3 779 893 074	3 879 408 885	1,03	1,03
500	22 440 811 654	22 872 440 605	22 915 639 213	4 746 658 850	4 738 936 680	1,00	1,00
1000	34 363 917 002	34 649 339 435	34 746 361 403	5 194 090 279	5 198 309 654	1,00	1,00

18. táblázat: Adott éves gyakoriságú (1. oszlop)  $\mu=10$  és  $\sigma=3$  paraméterű veszteségeloszláshoz tartozó aggregált eloszlás VaR-jának számítása az igazi paraméterekkel, valamint az (2)-es pontban megfogalmazott feltételek mellett számolt VaR értékek várható értéke és szórása, valamint ezek páronkénti aránya.

A visszamérés kapcsán megállapíthatjuk, hogy kisebb gyakoriságok ( $\lambda=[10, 20]$ ) esetén akár 10 százalékos várható érték - és szórásbeli eltérés is adódhat a szimulációs vizsgálathoz képest, ugyanakkor ez nem annyira durva, hogy ne használhatnánk modellünket következtetések levonására a tőkét illetően. Másrészt ez az eltérés szimulációs hibát is tartalmaz, hiszen mindössze csupán 10.000 kísérletből állapítottuk meg  $VaR_1$  és  $VaR_2$  jellemzőit.



20. ábra  $\text{VaR}_0$  várható értékének és szórásának a hibája a gyakoriság függvényében. Az (1)-es pontban megfogalmazott feltételek mellett számolt  $\text{VaR}_1$  értékek átlagához és szórásához viszonyítjuk  $\text{VaR}_0$  várható értékét és szórását. Az adatokat az eloszlásonkénti 13. táblázat - 16. táblázat utolsó két oszlopa tartalmazza.

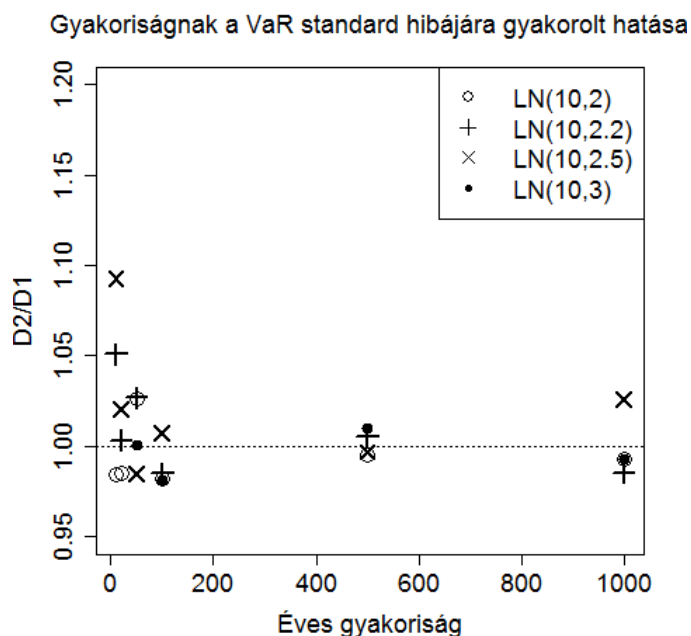


21. ábra  $\text{VaR}_0$  várható értékének és szórásának a hibája a gyakoriság függvényében. A (2)-es pontban megfogalmazott feltételek mellett számolt  $\text{VaR}_2$  értékek átlagához és szórásához viszonyítjuk  $\text{VaR}_0$  várható értékét és szórását. Az adatokat részben az eloszlásonkénti 17. táblázat - 18. táblázat utolsó két oszlopa tartalmazza.

Következtetések a modelltől és a visszamérésből

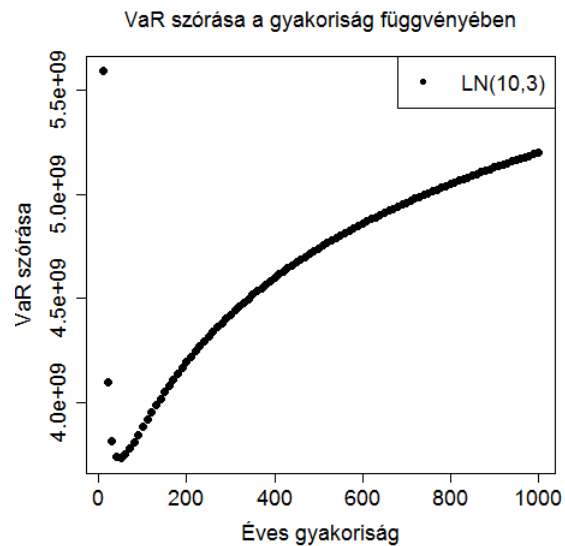
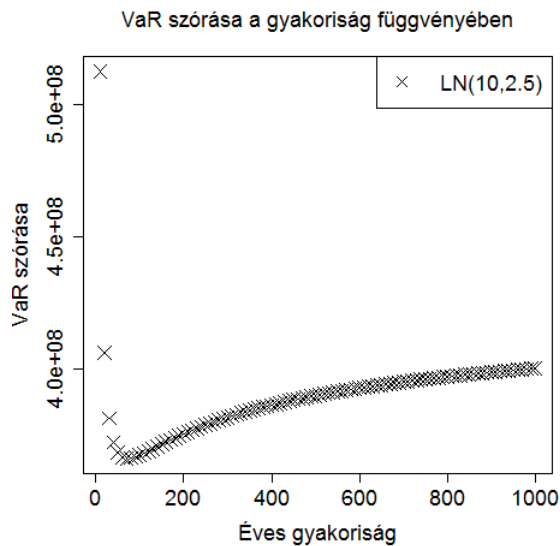
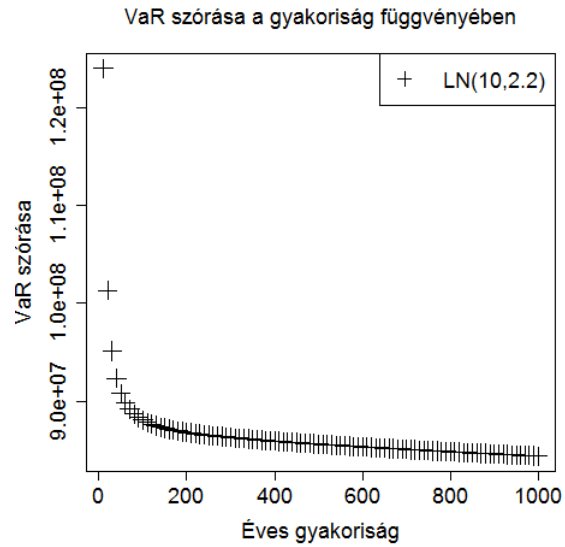
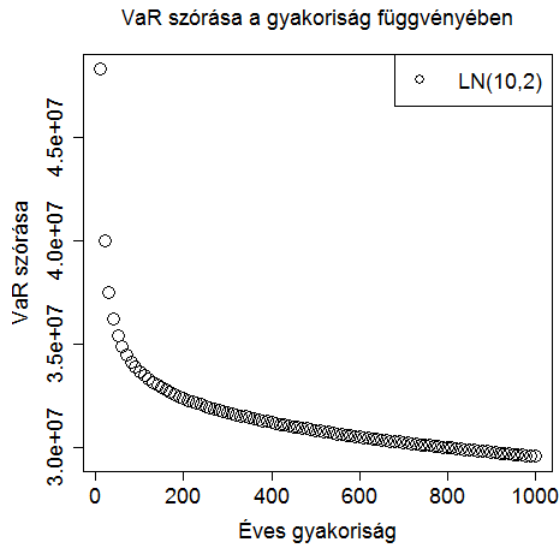
A következtetéseket lognormális veszteségeloszlás használata és Single Loss Approximation alkalmazása mellett hoztuk meg. A kapott eredmények tekintetében az további vizsgálatot igényel, hogy azok LDA környezetben és más eloszlások esetén is fennállnak-e?

1. A gyakoriság becslési hibája csak kismértékben befolyásolja a tőkét. Ez tulajdonképpen a szimulációs eredményekből látszik, hiszen az igazi (1-es modell) és a becült (2-es modell) gyakoriság paraméterek ( $\lambda$ ) melletti VaR szórások alig különböznek. Ez összhangban van a nemzetközi eredményekkel, miszerint a gyakoriság kevésbé fajsúlyos kérdés a tőke becslése során. (Aue et. al. [2006], Basel Committee on Banking Supervision [2008])



22. ábra: Az (1) és (2) szimulációs modellekből számolt  $VaR_2$  és  $VaR_1$  értékekhez tartozó  $D_2$  és  $D_1$  szórások aránya.

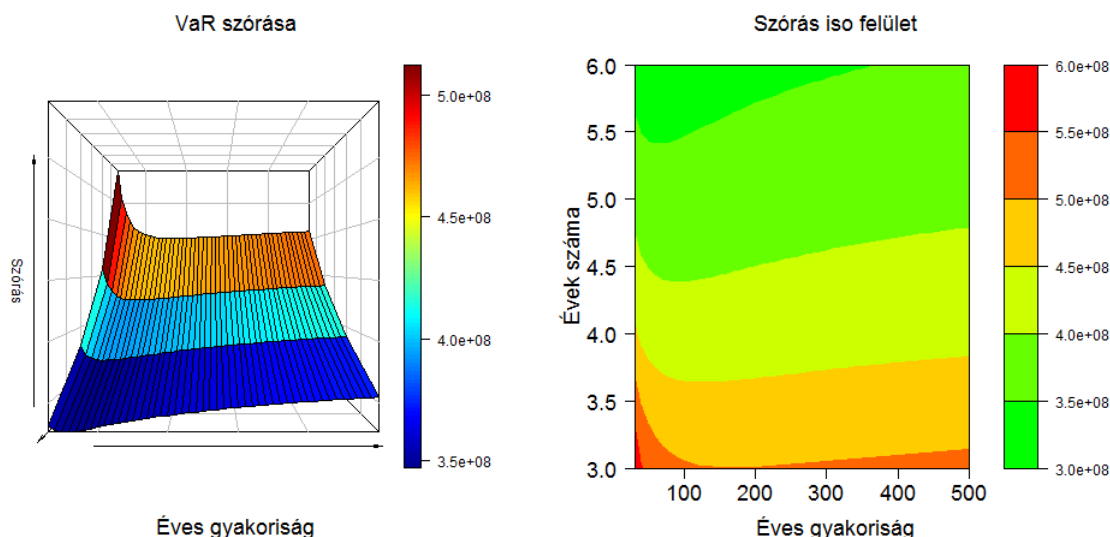
2. A becslés hibája (VaR szórása), akár a szimulációs, akár a levezetés eredményét tekintjük kevésbé vastag szélű ( $\sigma$  kicsi) esetben monoton csökkenő a gyakoriság függvényében, míg növekvő  $\sigma$  mellett a kezdeti csökkenés megtorpan és  $\lambda$  magasabb értékei mellett újra emelkedni kezd.



23. ábra: Poisson – Lognormális veszteséggeneráló folyamatból származó tőkekövetelmény sztenderd hibája a várható gyakoriság függvényében. A mintaelemszám ( $n$ ) az éves gyakoriság ötszöröse, azaz azt feltételezzük, hogy a bank 5 éves idősről becsül.

A jelenséget jobban átláthatjuk, ha a VaR szórását az éves gyakoriság és a becsléshez használt évek számának függvényében, valamint a rögzített gyakoriság mellett a mintaelemszám függvényében vizsgáljuk. Előbbi kapcsán a 24. ábra és az ahhoz kapcsolódó iso felszín nyújt segítséget.





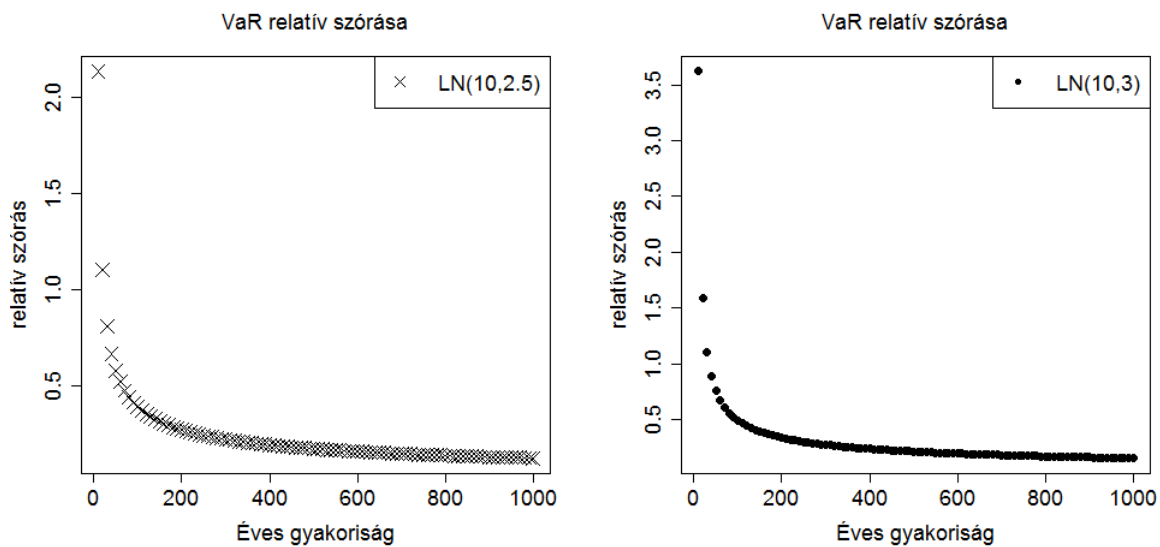
24. ábra: A bal oldali ábra (a) az éves gyakoriság (x) és az évek számának (y) függvényében mutatja a  $VaR_0$  szórását (z), míg a jobb oldali ábra (b) az (a) ábra iso szórás felületét mutatja. <sup>27</sup>

A 23. ábra érdekes helyzetet teremt a tőke becslése kapcsán. Nem csak az számít, hogy mennyi adattal rendelkezik egy intézmény, hanem az is, hogy mennyire vastag szélű a veszteséggeneráló folyamat. A kettő összjátéka határozza meg, hogy mennyire pontosan tudja az intézmény megbecsülni a működési kockázati tőkekövetelményét. Mint azt már korábban említettük két hatás feszül egymásnak. Egyrészt a növekvő gyakoriság kedvez a becslésnek, hiszen nagyobb mintából a veszteség eloszlást kisebb hibával tudjuk becsülni, másrészt viszont a nagyobb gyakoriság egyre távolabbi kvantilis extrapolálását teszi szükségessé. Úgy tűnik, hogy egy kevésbé kockázatos folyamat (kis  $\sigma$ ) esetén a mintanagyság növekedés ellensúlyozni tudja az extrapolációból adódó bizonytalanságot, addig vastag szélű folyamatok esetében ez már nem teljesül. Ebből adódóan paradox módon a mintanagyság növekedése a tőke becslésének nagyobb bizonytalanságával párosul vastag szélű folyamatok esetében!

<sup>27</sup> Az egyedi veszteséeloszlás lognormális  $\mu=10$  és  $\sigma=2,5$  paraméterekkel. Az (a) ábrát pontosan szemből látjuk (azaz az y tengely bentről kifelé mutat), az évek száma tőlünk legtávolabb 1. A jobb oldali ábra (b) az (a) ábra iso szórás felületét mutatja, azzal a különbséggel, hogy az y tengely szerinti évek száma tőlünk távolodva növekszik. Ez azonban nem változtat azon a jól látható jelenségen, hogy az éves gyakoriság függvényében vizsgálva  $VaR_0$  szórását az kezdetben csökken, majd emelkedni kezd. A 23. ábra ugyanezen paraméter melletti tagja éppen a 24(a) ábra azon speciális metszetét mutatta be, amikor az évek száma megegyezik a szabályozó által előírt értékkel, azaz  $y=5$ . A 24 ábráinak másik szembevető és egyben várt jellemzője, hogy egyedi veszteséeloszlás paramétereit minél több év adatából becsüljük, annak kisebb  $VaR_0$  szórása.

Megjegyezzük, hogy ez a jelenség LDA környezetben is fennáll, így nem tekinthető sem a modellezési környezet, sem az SLA sajátosságának!

Természetesen a kép ennél árnyaltabb, hiszen a növekvő szórás még mindig lehet elhanyagolható a tőkéhez képest, így érdemes a relatív szórásra is egy pillantást vetni. Relatív szórás alatt értsük most a szórásnak a tőkéhez (SLA-val számolva) viszonyított arányát, a megszokott várható értékkel szemben. Ezeket csak a  $\sigma=2,5$  és  $\sigma=3$  paraméterekre mutatjuk meg, hiszen láttuk, hogy a kisebb  $\sigma$ -jú eloszlások esetében maga a szórás is monoton csökkenő, így ott a relatív szórás monoton csökkenő lesz.

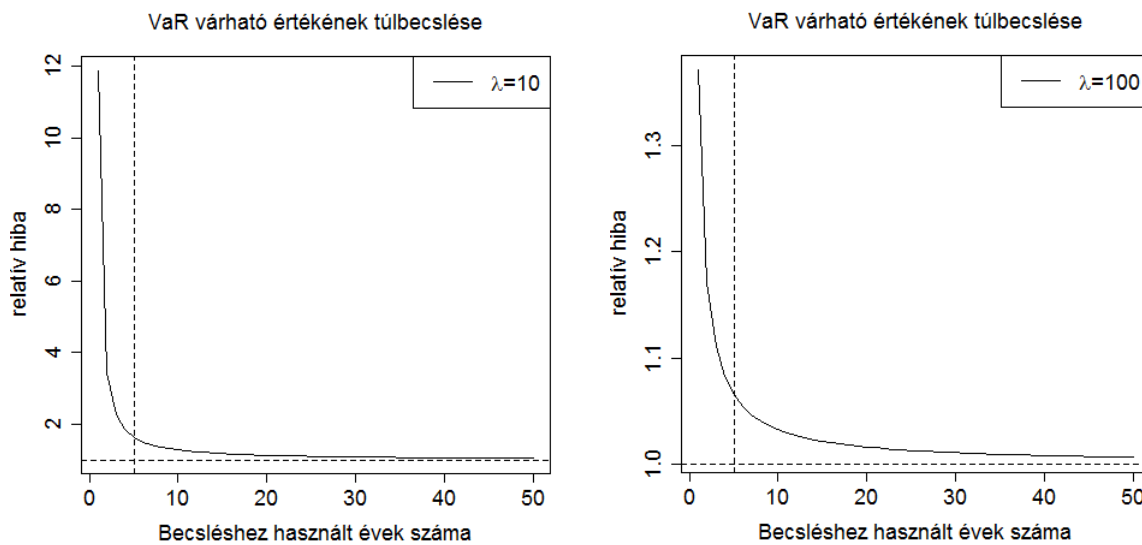


25. ábra: Relatív szórás alakulása  $\sigma=2,5$  és  $\sigma=3$  esetekben. Korábban láthattuk (23. ábra), hogy ezen esetekben a VaR szórása kis gyakoriság mellett csökken, majd újból nőni kezd. Amint az az ábrákon látható a relatív szórás már monoton csökkenő.

Megnyugtató a 25. ábra eredményei alapján, hogy a relatív szórás már monoton csökkenő. Ez azt jelenti, hogy a több veszteség eseménnyel rendelkező bankok relatíve mégiscsak jobb helyzetben vannak - eloszlástól függetlenül, hiszen a tőke nagyságrendjéhez viszonyított hiba monoton csökken. Tehát az extrapolációból adódó hátrányokat kompenzálja a növekvő mintaelemszám!

3.  $VaR_0$  várható értéke és szórása egyaránt függ a mintaelemszámtól ( $n$ ). Végtelen nagy minta esetén (limeszben)  $VaR_0$  várható értéke a becslési hibától mentes Single Loss Approximation eredményével egyezik meg. A mintaelemszám annál nagyobb (fix  $\lambda$

mellett), minél hosszabb ideje gyűjti a bank a veszteségadatokat. Ez persze nem mindig tehető meg, hiszen a minta nagysága mellett a veszteség generáló folyamat időbeli stabilitása is szükséges. Így a kettő összjátékán múlik, hogy a hitelintézet hány évnnyi adatból határozza meg a tőkét. Emellett még egyéb tényezők is szerepet kaphatnak, mint pl. hogy fejlett modell bevezetésekor a szabályozó 3 évnnyi időszorral is beérni, amit fokozatosan legalább 5 évre kell növelni. A következő ábrák rávilágítanak arra, hogy lambda kis és nagy értéke mellett az évek múlása mennyivel járul hozzá a pontosabb tőke meghatározáshoz.

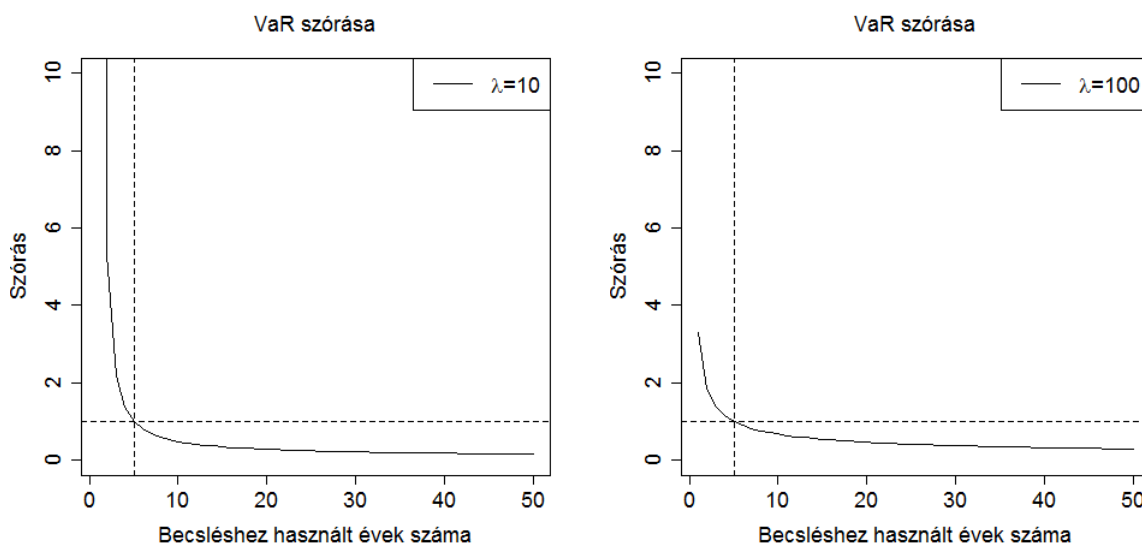


26. ábra: VaR<sub>0</sub> várható értékének és az igazi paraméterekkel számolt VaR (SLA) aránya a becsléshez használt mintaelemszám (avagy a felhasznált évek) függvényében.<sup>28</sup>

A 26. ábra jól rávilágít arra, hogy a kevés veszteséggel rendelkező intézmények várhatóan sokkal drasztikusabban becslik túl a működési kockázati tőkéjüket, mint a sok veszteséggel rendelkezők. Ez is azt a logikát erősíti meg, hogy az extrapolációs probléma hatása nem egyforma az egyes intézmények esetében. Minél kevesebb adattal rendelkezik egy intézmény, várhatóan annál inkább túlbecsli a működési kockázati tőkéjét.

Amint azt már korábban is láttuk, a becsléshez használt évek száma természetesen a VaR szórására is hatással van, itt is hozunk egy példát, hogy lássuk a kis és nagy gyakoriságú események közötti különbséget.

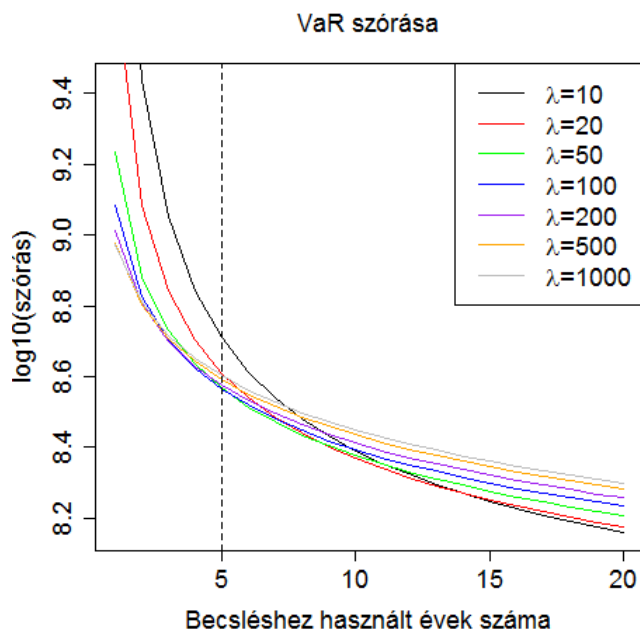
<sup>28</sup> Az egyedi veszteségeloszlás lognormális  $\mu=10$ ,  $\sigma=2,5$  paraméterekkel. A vízszintes szaggatott vonalat  $y=1$ -nél helyeztük el, hiszen límeszben a két érték aránya 1. Feltételeztük, hogy minden évben  $\lambda$ -val bővült a minta. A függőleges vonalat az 5. évnél húztuk meg, hiszen a bankok tipikusan 5 évnnyi mintából becslik az adataikat.



27. ábra:  $VaR_0$  szórásának és az 5 éves idősorból számolt  $VaR_0$  szórásának aránya a becsléshez használt mintaelemszám (avagy a felhasznált évek) függvényében.<sup>29</sup>

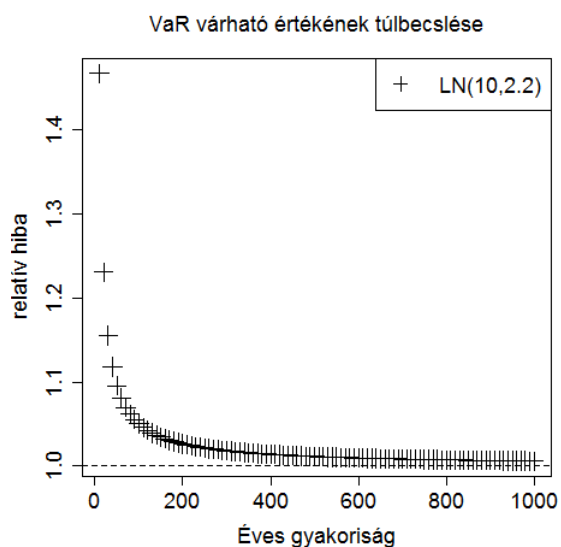
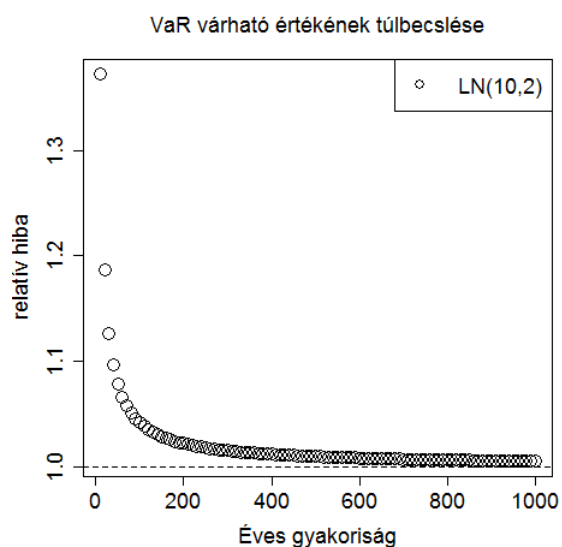
Szórás esetében az 5 évnyi adatból számított szóráshoz viszonyítunk (27. ábra), hiszen a különböző gyakoriságokhoz kapcsolódó VaR szórás értékek nagyságrendekkel különbözhetnek. Itt azt a következtetést vonhatjuk le, hogy rövid idősor (pl. a fejlett modell bevezetésekor lehetséges 3 év) esetén a VaR szórása többszöröse lehet az 5 éves idősorhoz tartozó szórásnak, amit a kis gyakoriság csak tovább fokoz. Ezek után már azt is látjuk, hogy miért lehet jelen a korábbi paradoxon, azaz hogy  $VaR_0$  szórása kezdetben csökkenő, majd növekedni kezd a gyakoriság függvényében. Az becsléshez használt évek növekedése csökkenti a becslési hibát, ugyanakkor a becslés dinamikája eltérő. Láthattuk (27. ábra) pl. hogy minél kisebb az éves gyakoriság, annál nagyobb a néhány évből becsült tőke hibája, ugyanakkor ez a hiba dinamikusan csökken az évek múlásával. Nagyobb gyakoriságú esemény becslési hibája pár évből történő becslés esetén sem annyira rossz, mint egy kis gyakoriságú eseményé, ugyanakkor a becslés hibája lassabban csökken az évek múlásával. A különböző indulási szintek (1 évből becsült VaR szórása) és az eltérő dinamika okozza a korábban vázolt paradoxont (lásd lenti 28. ábra). Megjegyezzük, hogy ha pl. 20 éves idősor állna a bankok rendelkezésére, akkor a szórás monoton növekvő lenne a gyakoriság függvényében, lásd 28. ábra.

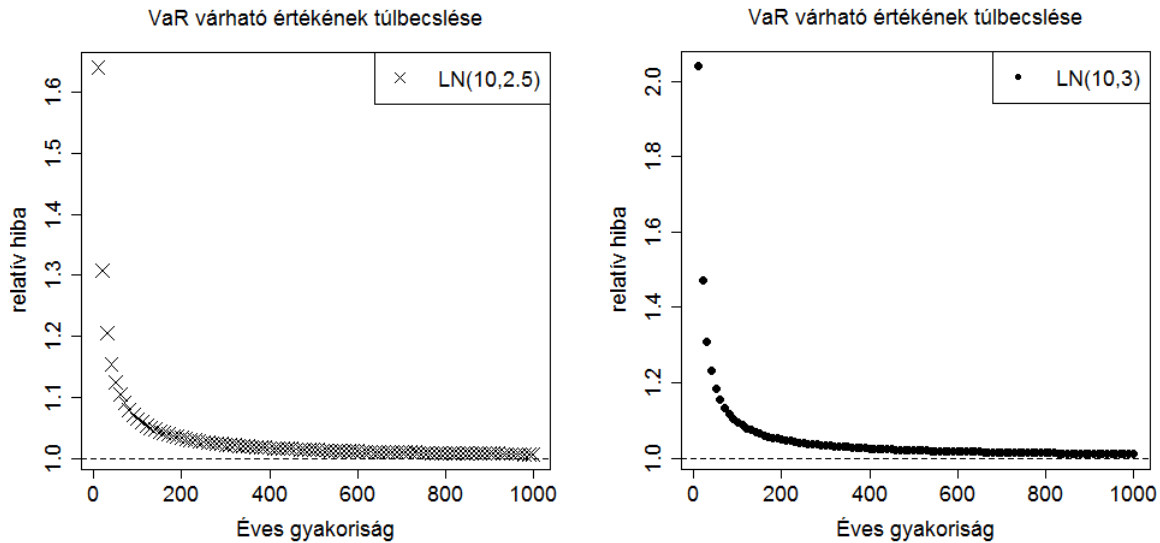
<sup>29</sup> Az egyedi veszteségeloszlás lognormális  $\mu=10$ ,  $\sigma=2,5$  paraméterekkel. A vízszintes szaggatott vonalat  $y=1$ -nél helyeztük el viszonyításképpen. Feltételeztük, hogy minden évben  $\lambda$ -val bővült a minta. A függőleges vonalat az 5. évnél húztuk meg, hiszen a bankok tipikusan 5 évnyi mintából becslik az adataikat.



28. ábra: VaR szórásának 10-es alapú logaritmusai különböző gyakoriságok mellett. A veszteség eloszlás paramétere:  $\mu=10, \sigma=2,5$ .

4. A kevés adattal rendelkező bankokat egy másik probléma is sújtja. A 13-18 táblázatokból jól látható, hogy a VaR várható értéke kis gyakoriság mellett (mind a szimulációs, mind az analitikus közelítés szerint) felülbecsli a tényleges, becslési hibától mentes VaR-t (29. ábra).

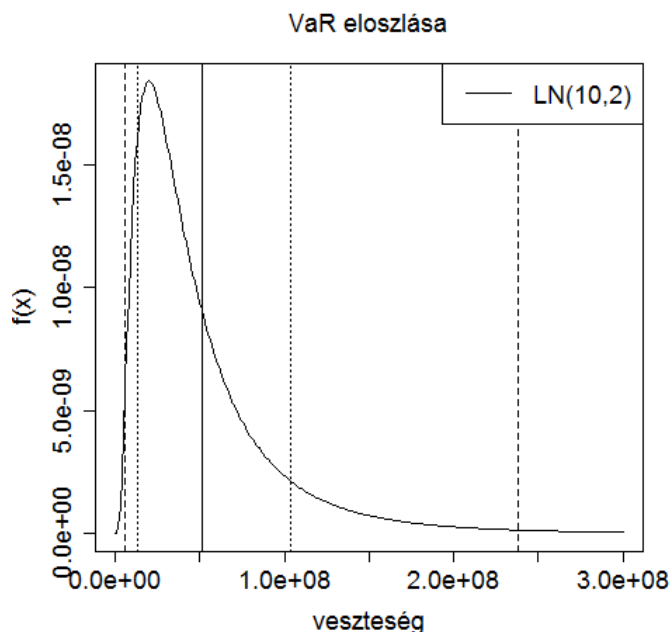




29. ábra: VaR<sub>0</sub> várható értékének és a Single Loss Approximation keretei között számolt VaR érték hányadosa. A minta nagyságát 5 évnnyinek feltételeztük ( $n=5\lambda$ ).

A 29. ábra alapján azt a következtetést vonhatjuk le, hogy a kockázatosabb folyamatok esetében szignifikánsabb a VaR túlbecslése.

5. VaR<sub>0</sub> eloszlása lognormális, így tehát az jobbra hosszan elnyúló, tehát a becslési hibák a túlzott tőkeképzésnek kedveznek. Ez összhangban van az empirikus megfigyelésekkel és a szimulációs kísérletekkel. (Baud et. al. [2002]). A 19. táblázat - 22. táblázat számos  $\lambda, \mu, \sigma$  paraméterre tartalmazza a várható tőkét, valamint azt a tartományt, ami 98 százalékos valószínűséggel tartalmazza azt, ahol  $Q_{01}$  és  $Q_{99}$  jelentik a tartomány határait. Emellett feltüntetjük még az  $R_{01} = Q_{01}/E(\text{VaR}_0)$  és  $R_{99} = Q_{99}/E(\text{VaR}_0)$  hányadosokat is, melyek a tőke alulbecsléséről, illetve túlbecsléséről adnak információt (1-1 százalék annak a valószínűsége, hogy  $Q_{01}$ -nél kevesebb és  $Q_{99}$ -nél több tőkét képezzünk).



30. ábra:  $\lambda=10$  gyakoriságú (tehát 5 év alatt átlagosan 50 elemű mintával rendelkező intézmény) és LN(10, 2) paraméterű egyedi veszteségeloszláshoz tartozó tőke eloszlása.<sup>30</sup>

lambda	mu	sigma	E(VaR <sub>0</sub> )	Q <sub>01</sub>	Q <sub>99</sub>	R <sub>01</sub>	R <sub>99</sub>
10	10	2	51.374.610	5.878.456	238.352.505	0,11	4,64
20	10	2	62.617.217	13.514.292	205.942.832	0,22	3,29
50	10	2	87.788.833	32.968.292	201.009.315	0,38	2,29
100	10	2	116.122.959	57.579.925	216.019.368	0,50	1,86
500	10	2	225.103.457	162.370.592	306.318.988	0,72	1,36
1000	10	2	297.750.852	235.237.374	373.187.514	0,79	1,25

19. táblázat: Várható tőke (E(VaR<sub>0</sub>)), és a tőkét 98 százalékos valószínűséggel tartalmazó tartalmazó intervallum (Q<sub>01</sub>-Q<sub>99</sub>) a gyakoriság és a veszteségeloszlás paramétereinek a függvényében.

lambda	mu	sigma	E(VaR <sub>0</sub> )	Q <sub>01</sub>	Q <sub>99</sub>	R <sub>01</sub>	R <sub>99</sub>
10	10	2,2	115.519.337	10.277.681	603.458.223	0,09	5,22
20	10	2,2	141.336.804	25.679.003	513.838.766	0,18	3,64
50	10	2,2	202.812.561	68.487.547	500.314.765	0,34	2,47
100	10	2,2	274.817.510	126.474.719	541.561.122	0,46	1,97

<sup>30</sup> A folytonos függőleges vonal a tőke várható értéke, míg a pontozott és szaggatott vonalakkal határolt tartományok 80 és 98 százalékos valószínűséggel tartalmazzák a megképzett tőkét. A várható érték nagyjából 50 millió forint, ugyanakkor 10% valószínűséggel több mint 100 millió forint és 10% valószínűséggel kevesebb mint 13 millió forint tőkét képez az intézmény. ezentúl 1% valószínűsége van annak, hogy a megképzett tőke meghaladja a 240 millió forintot.

500	10	2,2	567.244.725	395.606.693	795.237.459	0,70	1,40
1000	10	2,2	771.220.960	594.788.378	988.155.593	0,77	1,28

20. táblázat: Várható tőke ( $E(\text{VaR}_0)$ ), és a tőkét 98 százalékos valószínűséggel tartalmazó tartalmazó intervallum ( $Q_{01}$ - $Q_{99}$ ).

lambda	mu	sigma	$E(\text{VaR}_0)$	$Q_{01}$	$Q_{99}$	$R_{01}$	$R_{99}$
10	10	2,5	394 158 069	23 759 806	2 431 019 513	0,06	6,17
20	10	2,5	482 379 031	67 259 778	2 025 103 130	0,14	4,20
50	10	2,5	714 179 806	205 061 607	1 964 644 713	0,29	2,75
100	10	2,5	1 002 054 798	411 720 666	2 149 708 929	0,41	2,15
500	10	2,5	2 269 888 244	1 504 520 168	3 326 451 485	0,66	1,47
1000	10	2,5	3 215 493 631	2 391 369 574	4 257 676 652	0,74	1,32

21. táblázat: Várható tőke ( $E(\text{VaR}_0)$ ), és a tőkét 98 százalékos valószínűséggel tartalmazó tartalmazó intervallum ( $Q_{01}$ - $Q_{99}$ ).

lambda	mu	sigma	$E(\text{VaR}_0)$	$Q_{01}$	$Q_{99}$	$R_{01}$	$R_{99}$
10	10	3	3 146 159 121	96 033 442	24 794 603 629	0,03	7,88
20	10	3	3 796 523 160	334 747 663	19 913 500 531	0,09	5,25
50	10	3	5 865 077 106	1 275 475 921	19 202 238 739	0,22	3,27
100	10	3	8 690 742 386	2 943 975 800	21 392 750 627	0,34	2,46
500	10	3	22 915 639 213	13 940 830 687	36 123 387 410	0,61	1,58
1000	10	3	34 746 361 403	24 310 118 478	48 575 608 250	0,70	1,40

22. táblázat: Várható tőke ( $E(\text{VaR}_0)$ ), és a tőkét 98 százalékos valószínűséggel tartalmazó intervallum ( $Q_{01}$ - $Q_{99}$ ).

A táblázatból látható, hogy a gyakoriság növekedése mellett (veszteségeloszlástól függetlenül) szűkülnek a konfidencia intervallumok, így tehát a tőkebecslés stabilabbá válik. Másrészt minél inkább vastag szélű az egyedi veszteségeloszlás, annál inkább táguló konfidencia intervallumokra számíthatunk (táblázatok ugyanazon lambdáihoz tartozó  $R_{01}$  -  $R_{99}$  intervallumok összevetése).

## 8. Külső adatok elemzése

(önálló elemzés)



A külső adatok kötelező elemét jelentik a működési kockázati modelleknek, miközben használatuk előnyt és hátrányt egyaránt jelent a hitelintézetek számára. A külső adatbázis előnyeit és hátrányait korábban már bemutattuk. Legnagyobb előnye, hogy mérsékli azt az extrapolációs problémát, mellyel a hitelintézet a tőke becslése kapcsán szembesül. Amint azt korábban említettük a hitelintézetek tipikusan (5 évből történő becslés esetén) 200-szoros extrapolációval kell számolniuk, ami csökkenthető, ha hosszabb idősort tud figyelembe venni. Tegyük fel, hogy van 10 bank, melyek kockázati profilja pontosan megegyezik a veszteségeloszlás tekintetében, és megosztják egymással a működési kockázati eseményeiket. Ez esetben 1-1 bank úgy érzékeli, mintha 50 évnyi minta állna a rendelkezésére (10 bank 5 évnyi adata) és így az extrapolációs probléma méretét csökkenteni tudja. Az új helyzetben a banknak már csak egy 20-szoros extrapolációval kell megbirkóznia. A gyakorlatban ez nem feltétlenül így valósul meg. A hitelintézet valószínűleg csak egy küszöbön túl (ami nem egyezik meg az ő belső adatgyűjtési küszöbével) vonja be a külső eseményeket az eloszlás szélének a modellezéséhez. Összességében tehát a becslési hiba csökkenése várható, de még mindig lényeges hibázási lehetőség mellett!

### *8.1. A HUNOR adatbázisról*

A Bankszövetség és a Nemzetközi Bankárképző 2013-ban rendelkezésünkre bocsátotta a legfrissebb adatokat<sup>31</sup>, ezeket használjuk fel a dolgozat során az elemzésekhez, és a következtetések levonásához. Az adatok a 2009.06.30-2012.06.30 hároméves időszakra állnak rendelkezésünkre. Ez az időszak több szempontból is praktikus számunkra. Egyrészt nem tartalmazza a HUNOR adatbázis első 1-2 évét, amikor még számos koordinációs, adatgyűjtési, definíciós probléma nehezítette az adatok jó minőségének a biztosítását, másrészt nem tartalmazza a legfrissebb időszakot sem, melyből adatok hiányozhatnak az egyes szereplők késői adatszolgáltatása miatt.

### *8.2. Adatok előkészítése*

---

<sup>31</sup> Az adatokból származó eredményeket egy jelenleg még nem végleges közös (Deák Marianna, Gáll József, Nagy Gábor) publikáció tartalmazza.

A HUNOR adatkonzorcium 50.000 forinttól tartalmaz veszteségeket, ahol alacsonyabb veszteségek is bekerültek, azokat a kezdeti előkészítés során kiszűrtük. Szintén eldobtuk az adatbázisból minden olyan eseményt, ami nem tisztán a működési kockázati esemény (pl. boundary with credit, stratégia, reputáció ...), kivéve a boundary with market risk típusú eseményeket. A piaci kockázatok kezelésének módszertana és eszköztára (ellentétben a hitelkockázattal) egyáltalán nem alkalmas a vegyes típusú események (egyszerre tartalmaz piaci és működési kockázati elemet) kockázati tartalmának megragadására, így azokat a piaci gyakorlatnak megfelelően mi is figyelembe vesszük az elemzés során. Összességében az adatok jó minőségűnek tűnnek köszönhetően a Nemzetközi Bankárképző Zrt. rendszeres adattisztítási tevékenységének.

### *8.3. HUNOR statisztikák*

A leíró és feltáró statisztikák készítése során nem törekszünk a teljességre, hanem elsődlegesen azokra a szempontokra koncentrálnak, melyek működési kockázatok természetének a megértéséhez, és az eloszlás szélének vizsgálatához fontosak. A statisztikákat csak esemény típusonként készítjük, a bázeli üzletág dimenzió vizsgálatát mellőzzük.<sup>32</sup> Megjegyezzük, hogy Szabolcsné [2009] is foglalkozott a HUNOR adatbázis elemzésével, melynek tárgya a bankrendszer szükséges tőkekövetelménye volt. Ezzel szemben mi arra fókuszálunk, hogy a HUNOR (azaz valós) adatokon keresztül jobban megismerjük a működési kockázatok természetét, és a tőkeképzéssel kapcsolatos problémákat. Érdeklődésünk fókuszában továbbra is a tőke stabilitása áll. Számos banki tőkeképzési módszertan használja fel a HUNOR adatokat oly módon, hogy az egyedi veszteségeloszlást több részre osztja, és az alacsonyabb, mérsékeltebb tartományban a belső, míg az eloszlás szélén a HUNOR adatokból becsül. Ez utóbbit természetesen nem lehet minden veszteségkategória esetén megtenni (lásd pl. belső csalás kategória), hiszen még egy adatkonzorcium esetében is előfordulnak olyan veszteségkategóriák, ahol a mintaelemszám elégtelen. A következő táblázat egy rövid áttekintőt nyújt néhány esemény típusonkénti leíró statisztikáról.

---

<sup>32</sup> A bázeli üzletágak százalékosan megosztva kerülnek be a HUNOR adatbázisba, így ott azzal a kérdéssel is kell foglalkozni, hogy mi számít egy veszteségnek, azt érdemes-e egy osztályba sorolni, vagy azokat külön-külön kell kezelni?

Esemény típus	Darab	Átlag	Szórás	Medián	Maximum	Ferdeség	Csúcsosság
Belső család	27	83 546 145	249 704 836	3 841 978	1 219 692 706	4	13
Külső család (kivéve kártya)	287	4 789 556	39 122 766	234 682	521 200 093	12	141
Munkáltatói gyakorlat	133	3 777 092	13 141 218	300 000	92 330 477	5	25
Ügyfelek, termékek	413	5 600 699	43 421 380	200 000	601 607 289	13	166
Tárgyi eszközök	1 465	410 115	1 115 079	210 625	24 612 000	16	315
Üzletmenet, rendszerhiba	202	1 533 127	9 881 658	175 764	135 850 710	13	166
Végrehajtás, folyamatok	2 344	2 089 053	26 598 156	170 011	930 667 000	30	979
Kártya család	2 205	287 331	952 580	131 000	27 312 000	16	353

23. táblázat: HUNOR leíró statisztikák.

Az 23. táblázat alapján láthatjuk, hogy az események számossága 27 és 2344 között van kategóriánként. Mivel az eloszlás szélét kívánjuk vizsgálni (és ott is szükség van egy elégséges mintaelemszámra), így valószínűleg csak néhány kategória lesz alkalmas erre. Az egyes kategóriákban az események szórása többszöröse azok átlagának, a kategóriánkénti minta ferdesége és csúcsossága jelentős. A maximum és átlag arány esetenként a több százat is meghaladja. Mindezek arra utalnak, hogy a kategóriánkénti eloszlások közül több is erősen vastag szélű (lásd elsődlegesen a csúcsosságot).

Elsőként megnézzük, hogy az igen népszerű (és hasonló problémáknál gyakran jól viselkedő) lognormális eloszlás mennyire illeszkedik jól az adatokra. Egyben azt is megvizsgáljuk a Szabolcsné [2009] által korábban publikált (2007-2008 –as adatokból becsült) paraméterektől mennyire térnek el a jelenlegi becslések. Az eloszlás illesztése során az eredeti mintát eltoljuk 0-ba, és az így kapott adatokra illesztjük az eloszlást. Többek között Baud et. al. [2002], Ergashev et. al. [2012] megmutatták, hogy tőkeképzés szempontjából nem mindegy (jelentős különbségek adódhatnak!) hogy feltételes eloszlást használunk-e, vagy a 0-ba eltolta adatokra illesztjük-e a feltétel nélküli eloszlást. Mivel korábban már rávilágítottunk arra, hogy a feltételes eloszlás illesztése során milyen nehézségek adódnak (MLE becsülés esetén), így az

eltolások módszert választottuk. Az illesztések során Maximum Likelihood becslést használunk, és a KS és AD teszt statisztikák mellett helyenként grafikus módon is meggyőződünk az illesztés jóságáról.

Esemény típus	$\mu$	$\sigma$	KS statisztika	KS P érték	AD statisztika	AD P érték
Belső csalás	15,1432	2,7453	0,0813	0,9877	0,1980	0,9911
Külső csalás (kivéve kártya)	12,3283	2,2209	0,0458	0,5834	0,7400	0,5262
Munkáltatói gyakorlat	12,6497	2,2109	0,1011	0,1316	0,9495	0,3847
Ügyfelek, termékek	12,0686	2,3828	0,0839	0,0060	3,3636	0,0180
Tárgyi eszközök	11,8641	1,4198	0,0453	0,0049	7,8950	0,0001
Üzletmenet, rendszerhiba	11,7870	1,8272	0,0678	0,3102	1,4390	0,1917
Végrehajtás, folyamatok	11,6593	2,0702	0,0573	0,0000	8,3307	0,0001
Kártya csalás	11,1041	1,6282	0,0876	0,0000	17,5457	0,0000

24. táblázat: HUNOR adatbázison becsült  $\mu$ ,  $\sigma$  paraméterű lognormális eloszlás a hozzá tartozó Kolgomorow Smirnov, valamint Anderson-Darling teszt statisztikákkal és P értékekkel.

Esemény típus	$\mu$	$\sigma$	KS statisztika	KS P érték	AD statisztika	AD P érték
Belső csalás	15,1432	2,7453	0,0813	0,9877	0,1980	0,9911
Külső csalás (kivéve kártya)	12,4202	2,1984	0,0499	0,5071	0,8060	0,4766
Munkáltatói gyakorlat	13,1017	2,1807	0,0823	0,4945	0,7514	0,5170
Ügyfelek, termékek	12,3576	2,4231	0,0701	0,1391	1,2048	0,2657
Tárgyi eszközök	11,9601	1,3373	0,0394	0,0279	4,6086	0,0044
Üzletmenet, rendszerhiba	11,8526	1,9136	0,0610	0,5288	0,9749	0,3706
Végrehajtás, folyamatok	11,9770	2,0380	0,0353	0,0309	2,3832	0,0571

Kártya család	11,4556	1,4542	0,0404	0,0114	4,0563	0,0082
---------------	---------	--------	--------	--------	--------	--------

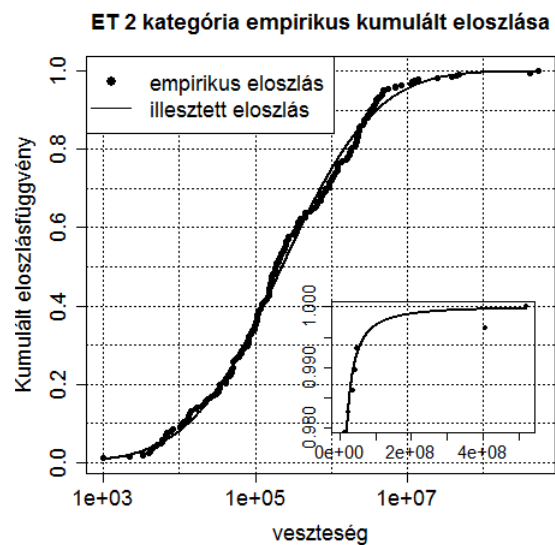
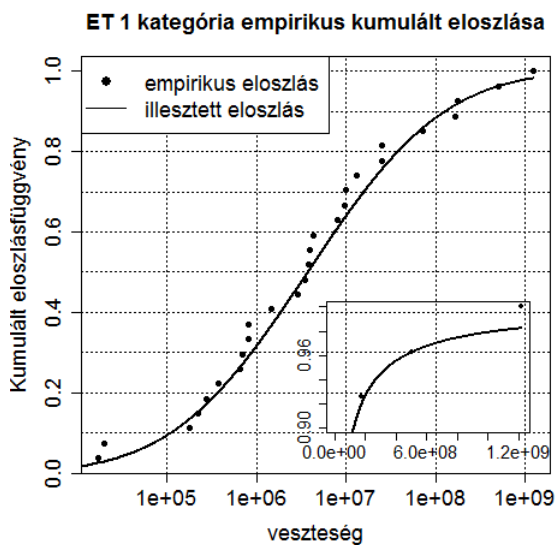
25. táblázat: Az azonos értékek kiszűrése után a HUNOR adatbázison becsült  $\mu$ ,  $\sigma$  paraméterű lognormális eloszlás a hozzá tartozó Kolmogorow Smirnov, valamint Anderson-Darling teszt statisztikákkal és P értékekkel.

A 24. táblázat adatiból látható, hogy megfelelő lognormális MLE illesztés csak a belső család eseménykategóriában érhető el. Elvégeztük momentumok módszerével és kvantilis illesztés módszerével is a becslést, de a különbségnek a tőkére gyakorolt hatását nem értékeltük ki, hiszen a becslés jósága érdemben nem változott. Megjegyezzük, hogy nagyméretű minták esetében már a minta kismértékű inhomogenitása is tönkretelheti a becslést. Ezen túlmenően az azonos mintaelemek is rontják a becslés jóságát MLE becslőfüggvény használata esetén. Ha megfigyeljük, hogy hol sikerültek a „legjobban” az illesztések, akkor éppen a legkisebb mintaelemszámú kategóriák köszönnek vissza. A minták számos esetben tartalmaznak azonos értékeket, ahol ennek köszönhetően ugrik az empirikus kumulált eloszlásfüggvény, ami szintén negatív hatással van GoF tesztekre. Ezen jelenség miatt a becslést elvégeztük oly módon is, hogy kiszűrtük a mintából az azonos értéket, feltételezve hogy azok egy másik – diszkrét eloszlásból származnak, és csak a megmaradt „folytonos” részre végeztük el az illesztést. Ilyen esetben az intézmény tőke modellje egy diszkrét és folytonos eloszlás keveréséből állna elő. Az azonos elemek kiszűrése után (25. táblázat) ugyan több kategóriában is javultak a GoF tesztek.

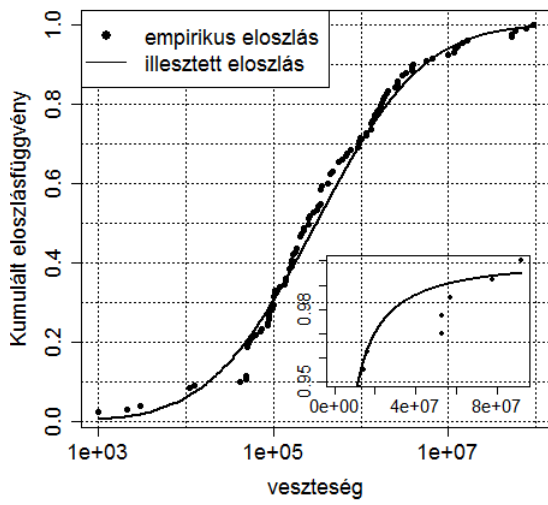
Esemény típus	$\mu_{2007-2008}$	$\sigma_{2007-2008}$	$\mu_{MLE}$	$\sigma_{MLE}$
Belső család	14,503	1,529	15,1432	2,7453
Külső család (kivéve kártya)	13,396	1,953	12,3283	2,2209
Munkáltatói gyakorlat	12,260	1,501	12,6497	2,2109
Ügyfelek, termékek	13,122	2,284	12,0686	2,3828
Tárgyi eszközök	12,077	1,326	11,8641	1,4198
Üzletmenet, rendszerhiba	11,887	1,575	11,7870	1,8272
Végrehajtás, folyamatok	12,334	2,285	11,6593	2,0702
Kártya család	10,905	1,176	11,1041	1,6282

26. táblázat: Hunor adatokra illesztett lognormális eloszlás paramétereinek az összehasonlítása a 2007-2008-as minta (Szabolcsné), és a Bankárképző Zrt. által rendelkezésünkre bocsájtott 2009.06.30-2012.06.30 időszakra vonatkozó minta felett.

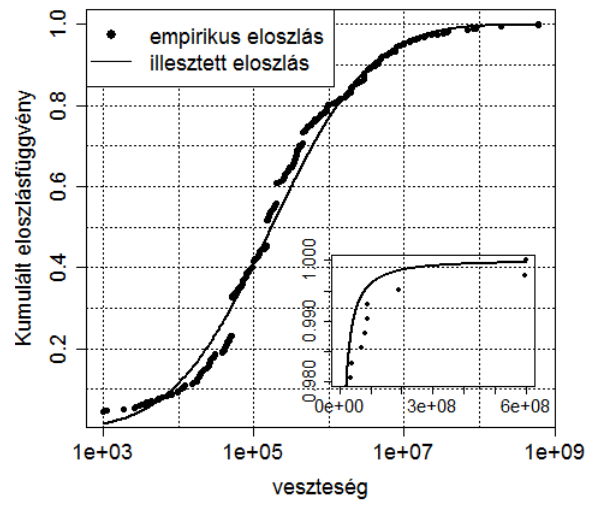
Annak ellenére, hogy a becslés összességében (eloszlás teste és széle együttesen) több kategóriában is megkérdőjelezhető, a HUNOR becslésnek több fontos mondanivalója is van. Nagyságrendileg leolvashatók azok a paraméterek, melyek mellett érdemes további szimulációkat, vizsgálatokat folytatni. Látható például, hogy a dolgozat során gyakran használt  $\mu=10$  és  $\sigma=[2, 2,5, 3]$  paraméterek közel állnak a realitáshoz, egy-egy intézmény kapcsán relevánsak lehetnek. Ezentúl a grafikus megjelenítés megmutatja számunkra, hogy az eloszlás szélén a lognormális eloszlás tipikusan alábecsli a kockázatokat (lásd az eseménytípusonkénti ábrák kinagyított része), így érdemes vastagabb szélű eloszlások irányába elmozdulni.



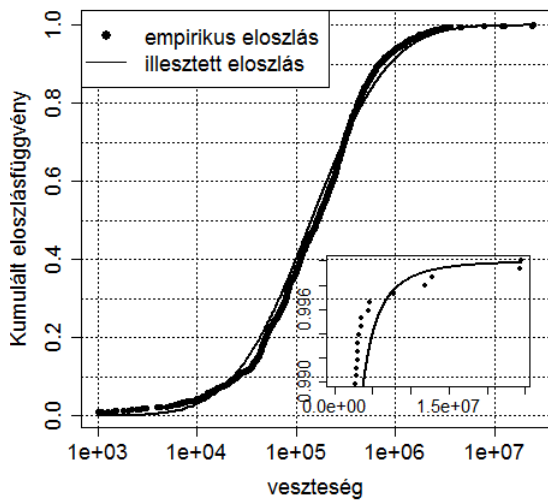
ET 3 kategória empirikus kumulált eloszlása



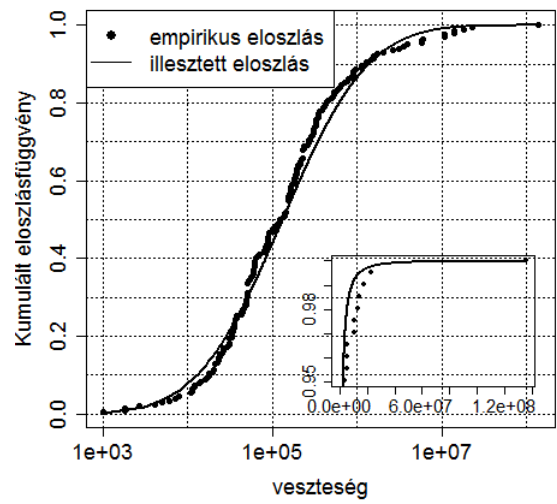
ET 4 kategória empirikus kumulált eloszlása



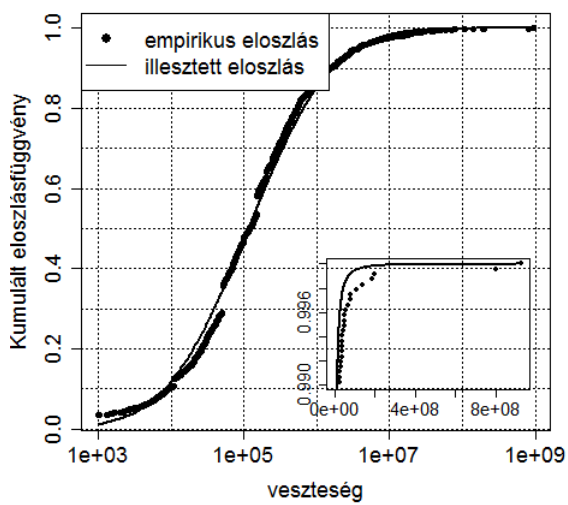
ET 5 kategória empirikus kumulált eloszlása



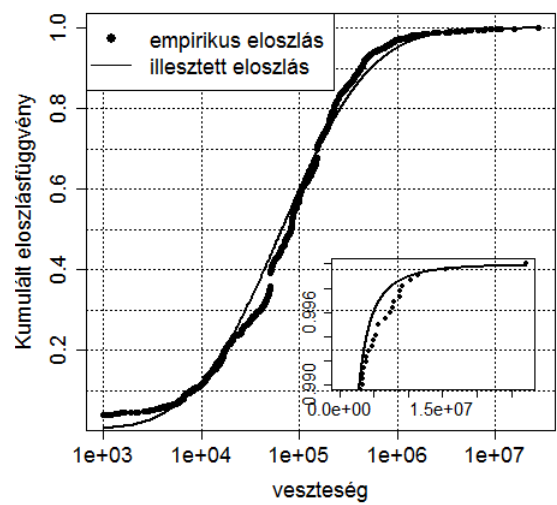
ET 6 kategória empirikus kumulált eloszlása



ET 7 kategória empirikus kumulált eloszlása



ET 8 kategória empirikus kumulált eloszlása



31. ábra: Lognormális eloszlás illesztése az egyes kategóriák veszteségeire.<sup>33</sup>

Érdemes megemlíteni, hogy az 1-es veszteségkategória illesztése igen jónak mondható, így hajlamosak lehetünk azt gondolni, hogy legalább egy kategóriában egyértelmű, hogy mennyi tőkét kell egy adott hitelintézetnek képeznie. De mielőtt megfedekezünk az extrapolációs probléma okozta instabilitásról (ami természetesen már csökkent a közös adatbázis használata által), rávilágítunk hogy egy ugyanolyan jó illesztés, ugyanazon eloszláscsaládon belül is akár 100%-os eltérést is okozhat a tőkében! Tegyük fel, hogy az adatbázist használó bank elfogadja, hogy a HUNOR alapján meghatározott veszteségeloszlás megegyezik a saját banki veszteségeloszlásával, ugyanakkor az események éves gyakorisága eltérő: mindössze 4 évente egyszer ( $\lambda=0,25$ ) következik be belső csalás az intézményben. A 26. táblázat tartalmazza a belső csalás kategória az MLE becslését, alább megmutatjuk, hogy kvantilis-kvantilis illesztéssel több olyan hasonlóan jó becslés is adható, melyek KS vagy AD értékei az MLE becsléssel esetenként megegyeznek (3 tizedjegyig).

Becslés típusa	Valószínűség1	Valószínűség2	$\mu$	$\sigma$	P érték KS	P érték AD	Tőke (SLA)
Maximum Likelihood	-	-	15,1432	2,7453	0,9877	0,9911	5.477.569.095
Kvantilis-kvantilis	0,3	0,84	15,0983	2,9008	0,9932	0,9912	7.910.241.101
Kvantilis-kvantilis	0,27	0,78	15,0492	2,5755	0,9876	0,9666	3.178.279.311

27. táblázat: Egymással azonosan jó becslések, melyek nagymértékben különböző tőkét eredményeznek.

A fenti táblázat kiemelt példáin túl még számos olyan jó becslést lehet adni csak a lognormális eloszlásra (tehát más eloszláscsaládot nem is vizsgálva), amely a fentiekől eltérő tőkét eredményez. Az általunk vizsgált esetek közül (kvantilis-kvantilis becslés, ahol az első valószínűség a [0,1:0,3], míg a második valószínűség a [0,7:0,9] intervallumban vehet fel értéket) azokat kiválasztva, ahol mind a KS-hez, mind az AD-hez tartozó p érték a 0,9-et

<sup>33</sup> Az ábra logaritmusos skálán készült, ugyanakkor az x tengelyen az eredeti veszteség értékek vannak feltüntetve. Az ábrák jobb alsó részén az eloszlás jobb széle van kinagyítva. (Az azonos mintaelemek nincsenek kiszűrve – habár elkészítettük, mivel az ábrák lényegében – az eloszlás szélét tekintve nem változtak a szűrés után.)



meghaladta a legkisebb becült SLA tőke 2.321.276.913 forint, míg a legnagyobb becült SLA tőke 14.791.545.574 forint volt. Ez egy 6-7 - szeres különbséget jelent a tőkékben, miközben a becslés eredménye (p értékek 0,9 fölött) tankönyvi példa lehetne!

#### 8.4. Extrém értékek elmélete

Az extrémek vizsgálatának egyik eszköztára<sup>34</sup> (Embrechts et. al. [2003]) a Peaks Over Threshold (POT) módszer, ami szerint az eloszlás széle egy bizonyos  $u$  küszöbön túl (limeszben) az általánosított Pareto eloszláshoz tart. Bár a név azt sugallja, hogy itt mindig vastag szélű eloszlásokról van szó, tudni kell, hogy az eloszlás szélének viselkedését egyetlen paraméter írja le ( $\xi$ , shape parameter, tail index), ami bizonyos érték mellett pontosan az exponenciális eloszlást adja vissza. A kétparaméteres GPD kumulált eloszlásfüggvénye a következő:

$$GPD_{\xi,\sigma}(x) = \begin{cases} 1 - \left(1 + \xi \frac{x}{\sigma}\right)^{-\frac{1}{\xi}} & \text{ha } \xi \neq 0 \\ 1 - e^{-\frac{x}{\sigma}} & \text{ha } \xi = 0, \end{cases} \quad (25)$$

ahol  $x \geq 0$  ha  $\xi \geq 0$ ;  $-\frac{\sigma}{\xi} \geq x \geq 0$  ha  $\xi < 0$ .

A kétparaméterű GDP eloszlás paramétereit a következőképpen interpretálhatjuk:

- $\xi$  foglalja magában az eloszlás szélével kapcsolatos információkat. A fenti képletből látható, hogy ha  $\xi = 0$ , akkor éppen az exponenciális eloszlást kapjuk vissza. Ha  $\xi > 0$ , akkor pedig a Pareto eloszláshoz jutunk, jelezve az adatok heavy tail viselkedésének a jellegét. Abban az esetben ha  $\xi < 0$ , az eloszlás felülről korlátos, így  $-\frac{\sigma}{\xi}$ -nél nagyobb eseményeknek 0 a valószínűsége. Megjegyezzük, hogy az exponenciális és Pareto eloszlások csak egy-egy speciális esetét jelentik azon eloszlásoknak, illetve eloszlás szél viselkedéseknek, amit a GPD eloszlás magában foglal. Az eloszlások szélének alakja ezen két speciális esetet meghaladóan egy szélesebb spektrumon mozog.

<sup>34</sup>

Lásd még GEV.

- $\sigma$  az eloszlás úgynevezett scale paramétere. A scale paraméter határozza meg az eloszlás skáláját, hogy mennyire koncentrált, vagy éppen kiterjedt (szétterülő) az eloszlás.

Legyen  $F_x(x)$  a működési kockázati események eloszlása egy adott kockázati osztályban, valamint jelölje  $F_u(y)$  ugyanezen kockázati osztály eseményeinek az eloszlását egy  $u$  küszöbön túl.<sup>35</sup> Ekkor  $F_u(y)$  az  $F_x(x)$  feltételes eloszlása, és érvényesül az alábbi összefüggés:

$$F_u(y) = P(X - u \leq y | X > u) = \frac{F_x(x) - F_x(u)}{1 - F_x(u)}, \quad (26)$$

ahol  $y = x - u$  és  $x - u > 0$ .

A Pikands Balkema-De Haan tétel állítása szerint (Embrechts et. al. [2003]) az eloszlások egy széles körére teljesül, hogy  $F_u(y)$  az általánosított Pareto eloszláshoz tart, amennyiben az  $u$  küszöbvel tartunk a végtelenbe.

$$\limsup_{u \rightarrow x_F} |F_u(y) - GPD_{\xi, \beta}(y)| = 0$$

$$GPD_{\xi, \beta}(y) = \begin{cases} 1 - \left(1 + \xi \frac{y}{\beta}\right)^{-\frac{1}{\xi}} & \text{ha } \xi \neq 0 \\ 1 - e^{-\frac{y}{\beta}} & \text{ha } \xi = 0, \end{cases} \quad (27)$$

ahol  $y = x - u$ ,  $\xi$  = shape paraméter, és  $\beta$  = scale paraméter.  $GPD_{\xi, \beta}(y)$ -t az angol nyelvű szakirodalom excess GPD -nek nevezi, ezzel is utalva arra, hogy  $y$  az  $x$  veszteség  $u$  küszöböt meghaladó részét (exceedances) jelenti.

### 8.5. Az eloszlás szélének a vizsgálata

Embrechts et. al. [2003] megmutatta, hogy a Mean Excess Function (MEF) alkalmas diagnosztikai eszköz az eloszlás szélének a vizsgálatához. A Mean Excess Function háromféle alakot ölthet:

- csökkenő trend (negatív meredekség), ami rövid szélű (short tailed) utal;
- nagyjából vízszintes, ami exponenciálisan lecsengő eloszlásokra jellemző;

<sup>35</sup> Megjegyezzük, hogy  $u$  nem feltétlenül egyezik meg az adatgyűjtési küszöbvel.

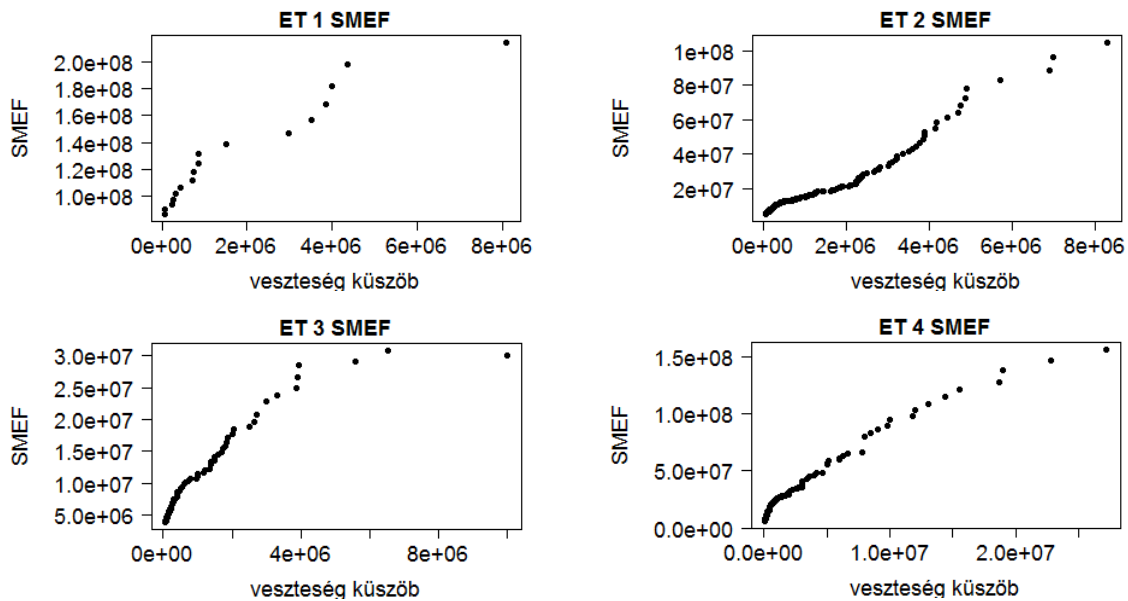
- emelkedő trend (pozitív meredekség) heavy tail eloszlások esetén.

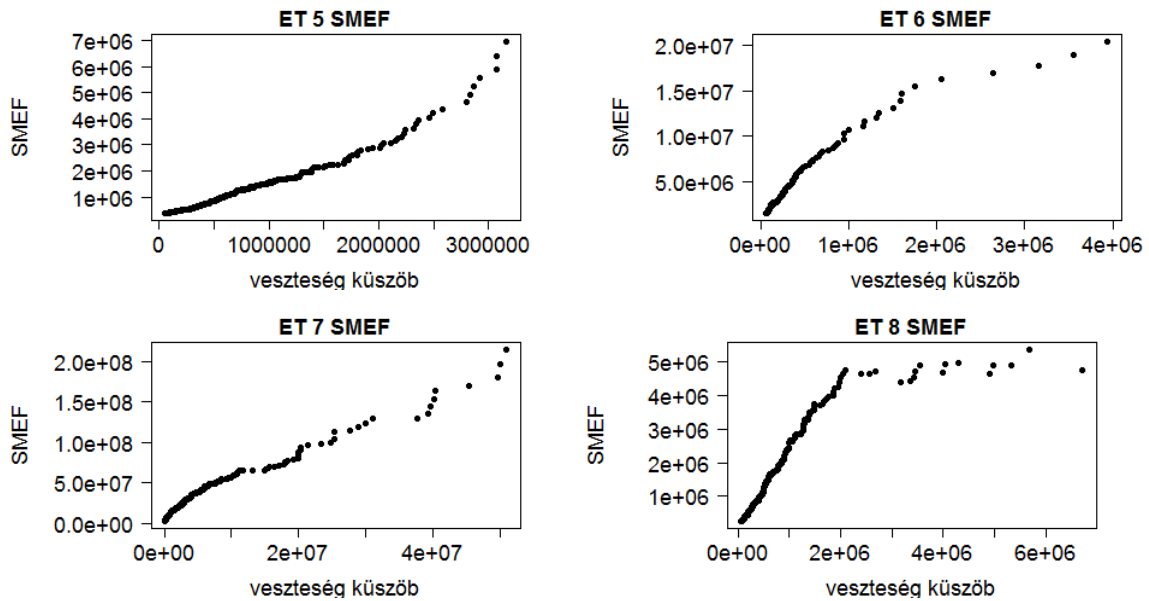
$$MEF(u) = E(X - u | X > u)$$

Ezen túlmenően a GPD eloszlás esetén a MEF lineáris függvénye az  $u$  küszöbnek. Ebből adódóan ez egy diagnosztikai eszközt is jelent, annak megállapítására, hogy milyen küszöböt érdemes feltételezni a GPD eloszlás becslése során. Tipikusan ahol a MEF lineárisra kezd válni  $u$  függvényében, azt érdemes modellezési küszöbnek választani. A MEF-en túl még egyéb diagnosztikai eszközök is léteznek az  $u$  megválasztására, lásd pl. Danielson et. al. [1997], Danielson et. al. [2001].

A Hunor adatbázis eseménytípusaira (kivéve belső csalás) elvégezzük a MEF becslését. Nem célunk az optimális küszöb megválasztása, sokkal inkább csak a MEF trendje érdekel minket az eloszlás szélének megismerése miatt. Az optimális küszöbvel azért nem foglalkozunk, mert a dolgozat során sokkal inkább foglalkoztat az a kérdés minket, hogy mekkora pontossággal határozhatóak meg egyáltalán a GPD eloszlás paraméterei és azon túl a tőke? Természetesen elvégezzük a paraméterek becslését a HUNOR adatok felett, hogy jobban értsük a tőkebecslés problémájának természetét, ugyanakkor a fő kérdés az, hogy mennyit is ér a becslés, mennyire támaszkodhatunk az eredményekre?

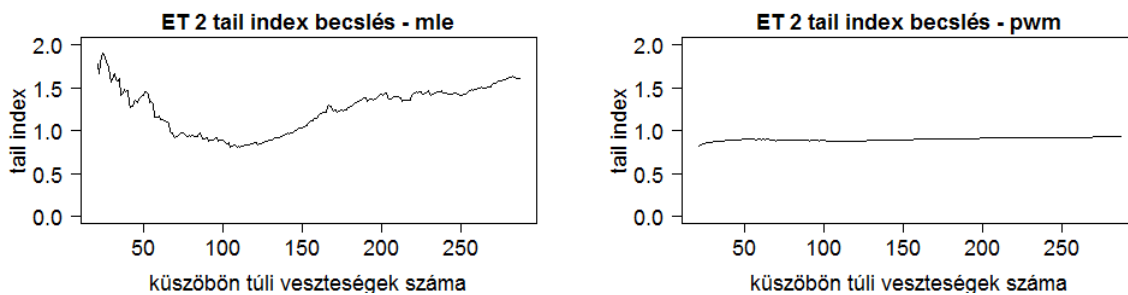
A SMEF (Sample MEF) ábrázolása során az utolsó 10 adatból (a monoton növekvő rendezett mintában) származó értéket már nem jelenítettük meg, mivel ott a kis mintaelemszám miatt annak becslése már igen instabil. A SMEF értékét az adatgyűjtési küszöbtől számítottuk.

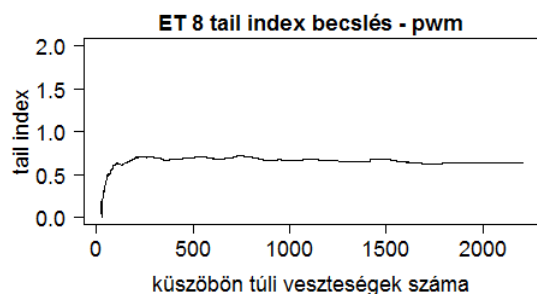
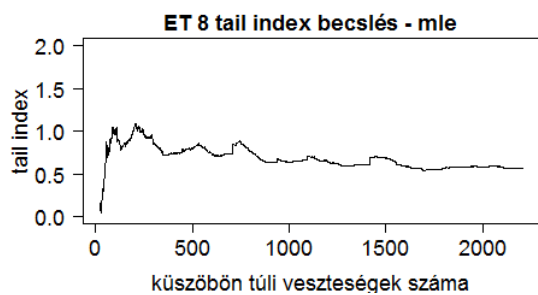
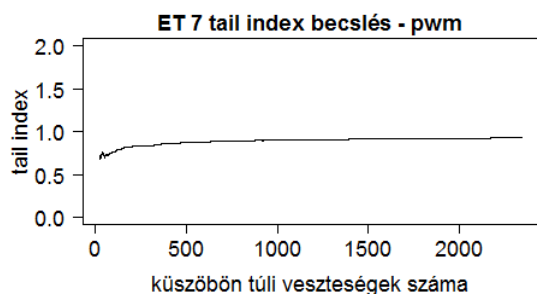
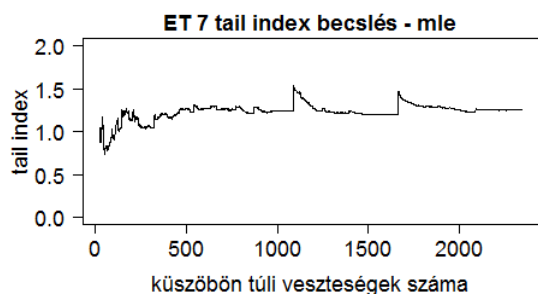
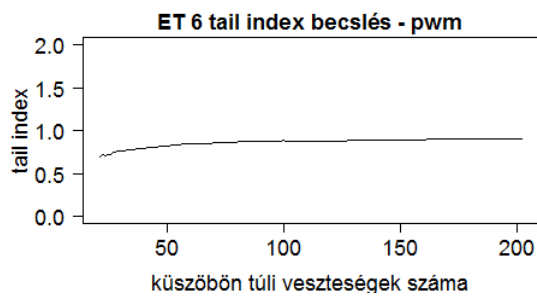
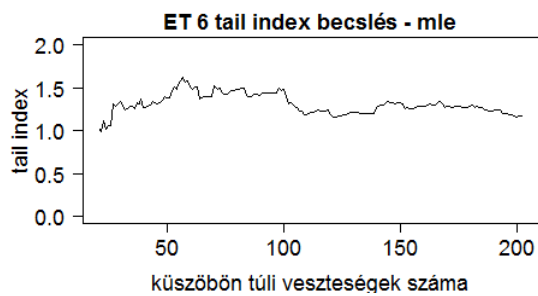
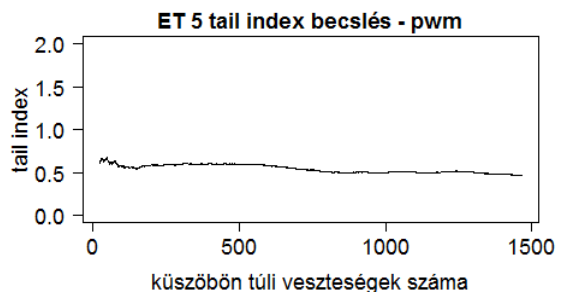
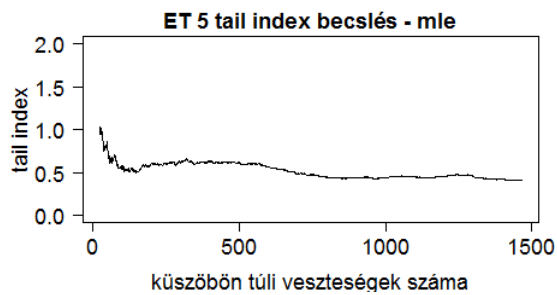
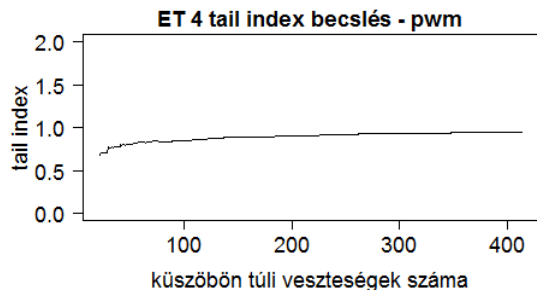
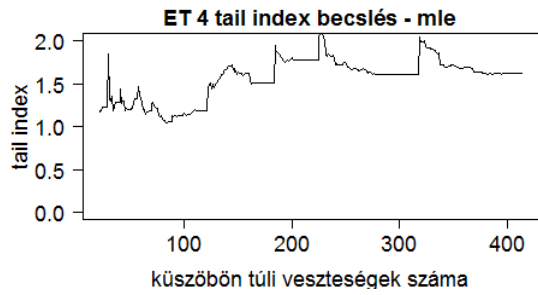
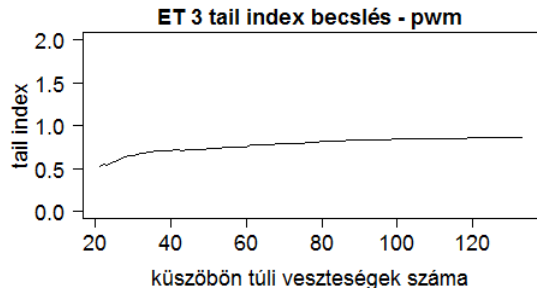
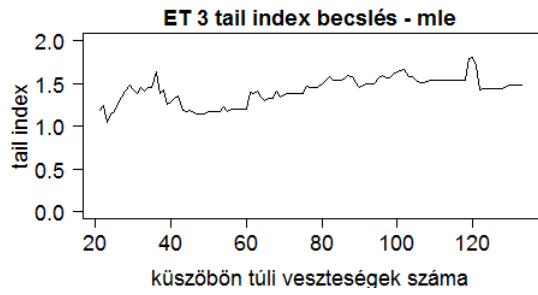




32. ábra: Sample Mean Excess (SMEF) az ET1-ET8 kategóriákra.

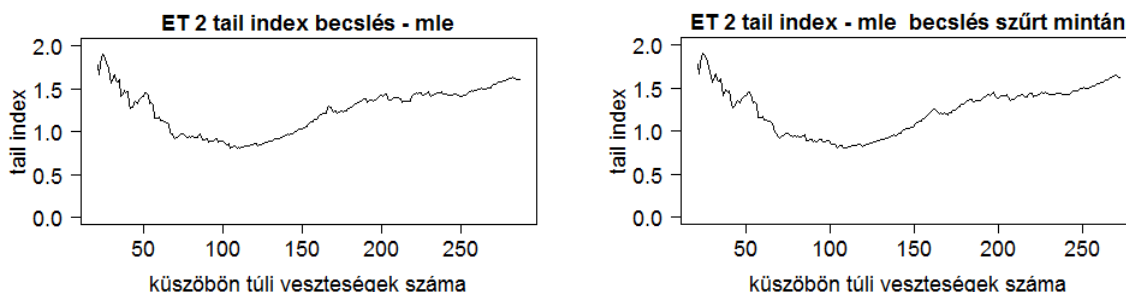
A 32. ábra alapján láthatjuk, hogy a Sample Mean Excess Function minden kategóriában pozitív trendet mutat, tehát a tail index pozitív értékére számítunk. Ez is megerősít minket abban, hogy a működési kockázati veszteségek vastag szélű eloszlásból származnak, így ilyen eloszláscsaládokat érdemes figyelembe venni. Az első kategória (ET1) esetében a GPD becslés nem kivitelezhető a kis mintaelemszám miatt, a többi kategóriában kísérletet teszünk a becslésre. ET1 kategória kivételével a veszteségek száma mindenhol meghaladja a 100-at, amit a szakirodalom (Embrechts et. al. [2003]) gyakran minimális mintaelemszámként említ GPD becslések esetén. Ugyanakkor mi a GPD paraméter becslését (maximum likelihood és súlyozott momentumok módszerével – PWM egyaránt) már egy 20 elemű mintától (top 20) kezdve elvégezzük, hogy lássuk a paraméterek viselkedését abban a tartományban is. A PWM becsléssel azért foglalkozunk, mert számos szerző (Cruz [2002], Chernobai et. al. [2006]) javasolja mint robusztus becslést a működési kockázati tőke meghatározása során.

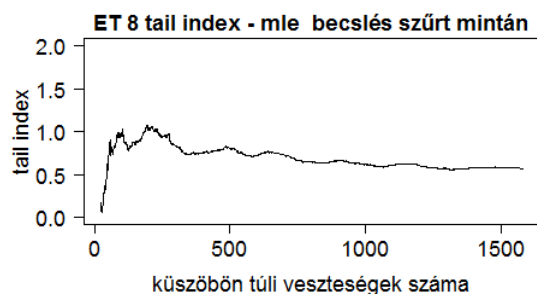
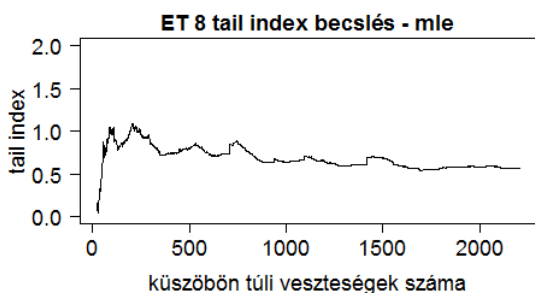
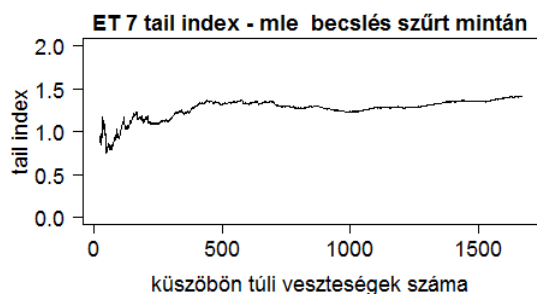
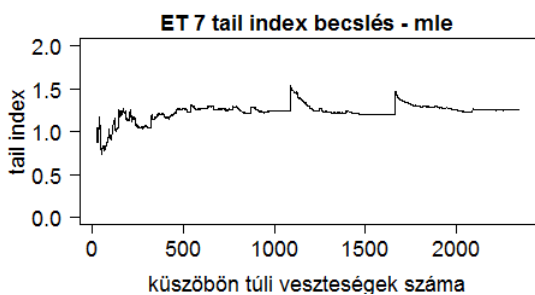
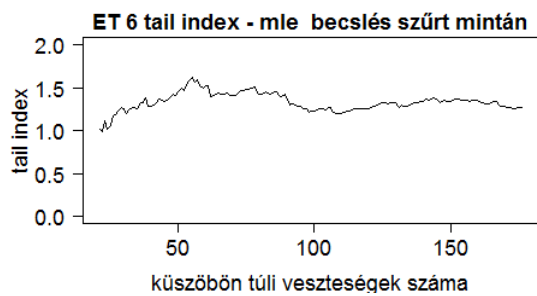
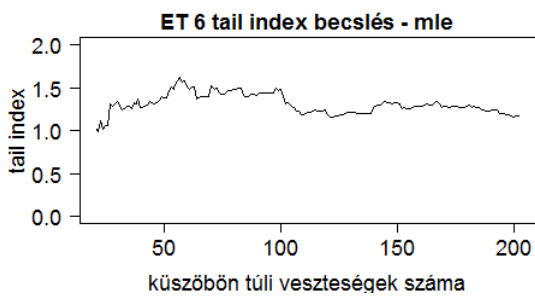
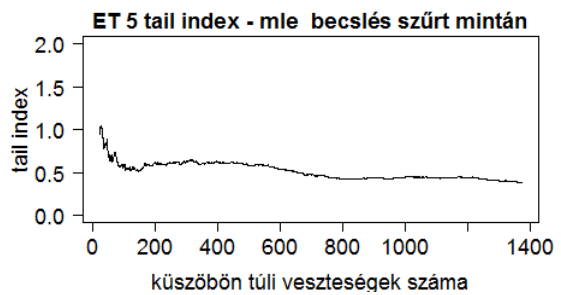
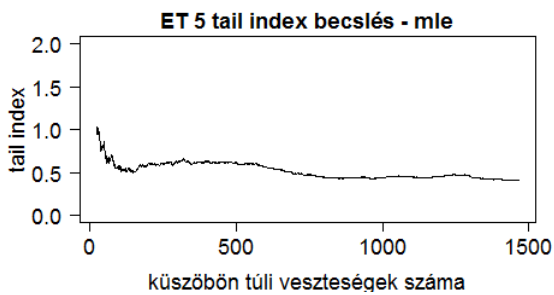
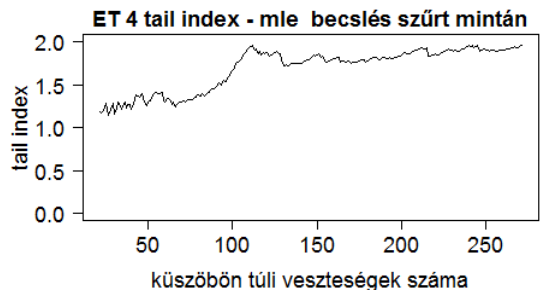
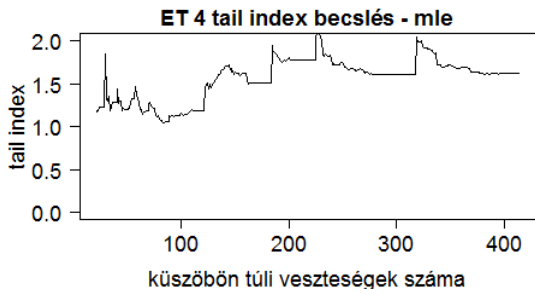
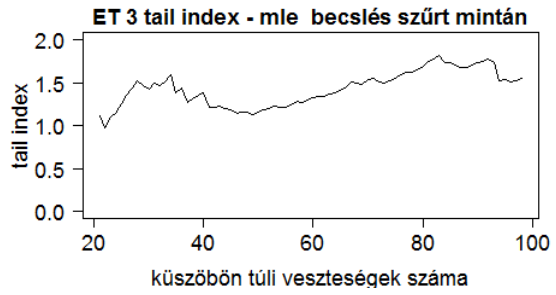
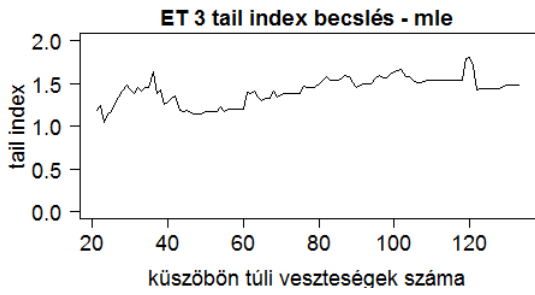




33. ábra: HUNOR adatbázisra MLE és PWM módszerekkel illesztett GPD eloszlás  $\xi$  paramétere.

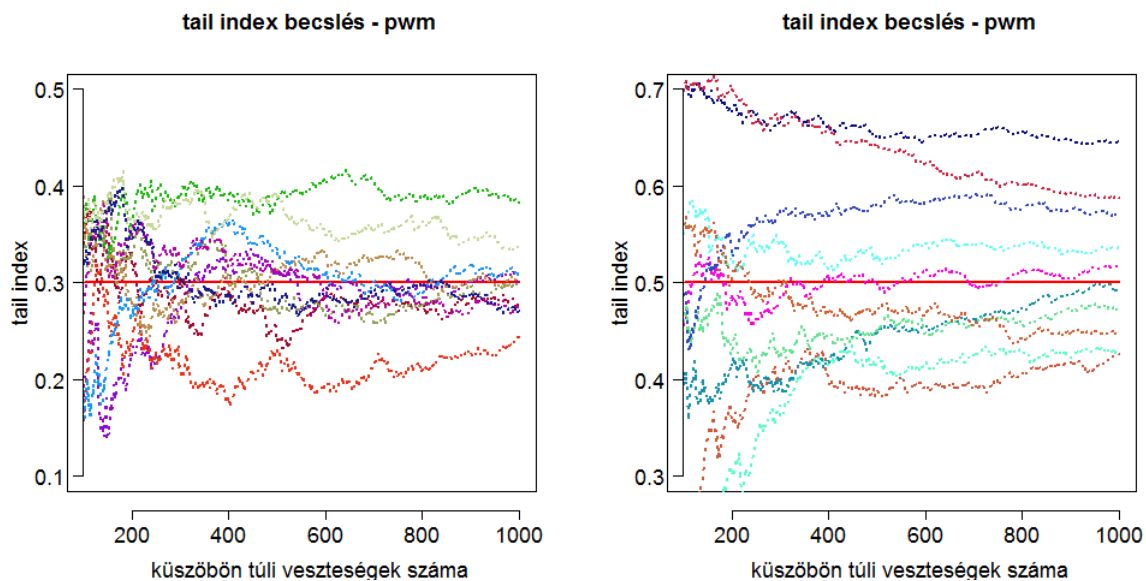
A HUNOR adatok GPD becslésén általánosságban látszik, hogy a PWM módszer egy sokkal simább, a küszöb függvényében stabil becslést eredményezett. Az MLE becslés esetében szembeűnő, hogy bizonyos esetben a becslés igen érzékeny a küszöb a változására, és a paraméter értéke ugrásszerűen megváltozik. Ez a jelenség figyelhető meg pl. a 33. ábra ET7 kategóriájában, ahol a küszöbön túli veszteségek száma 1100 és 1700 körűli és az MLE becslés fűrészfogszerűen változik. Ennek oka a már korábban, a lognormális eloszlás illesztése kapcsán is említett probléma. A minta bizonyos esetekben azonos értékeket tartalmaz, így az MLE becslés elromlik. Kérdés, hogy valóban arról van-e szó, hogy elromlik a becslés, vagy arról, hogy nem megfelelő eloszlást illesztettünk? Az empirikus mintában bizonyos értékek többször előfordulnak, így a hozzájuk tartozó relatív gyakoriságok megnövekednek, azaz a valószínűségek bizonyos pontokra koncentrálnak. Ez a tulajdonság nem jellemző a nevezetes eloszlásokra, így ezt a helyzetet mindenképp kezelni kell. Felmerülhet a koncentrált értékhez tartozó súly szétkenése az adott érték egy környezetében, de eljárhatunk úgy is, ahogy azt a lognormális illesztés esetében tettük. Az empirikus eloszlást egy olyan kevert eloszlásként fogjuk fel, ahol az egyik eloszlás egy diszkrét eloszlás, és pontosan azokat az értékeket tartalmazza, melyek többször fordulnak elő a mintában, míg a másik egy folytonos eloszlás. A diszkrét és folytonos eloszlások keverését pontosan az eredeti mintában hozzájuk tartozó realizációk valószínűségi ösztömege alapján valószínűsítjük meg a tőke modellezése során. Mivel a diszkrét eloszlás empirikusan adott, így a hangsúlyt a folytonos komponensre helyezzük, és újra elvégezzük a GPD eloszlás MLE becslését, miután kiszűrtük belőle a többszörösen előforduló értékeket, azaz pontosan fogalmazva, azokból csak egy realizációt hagyunk a mintában.



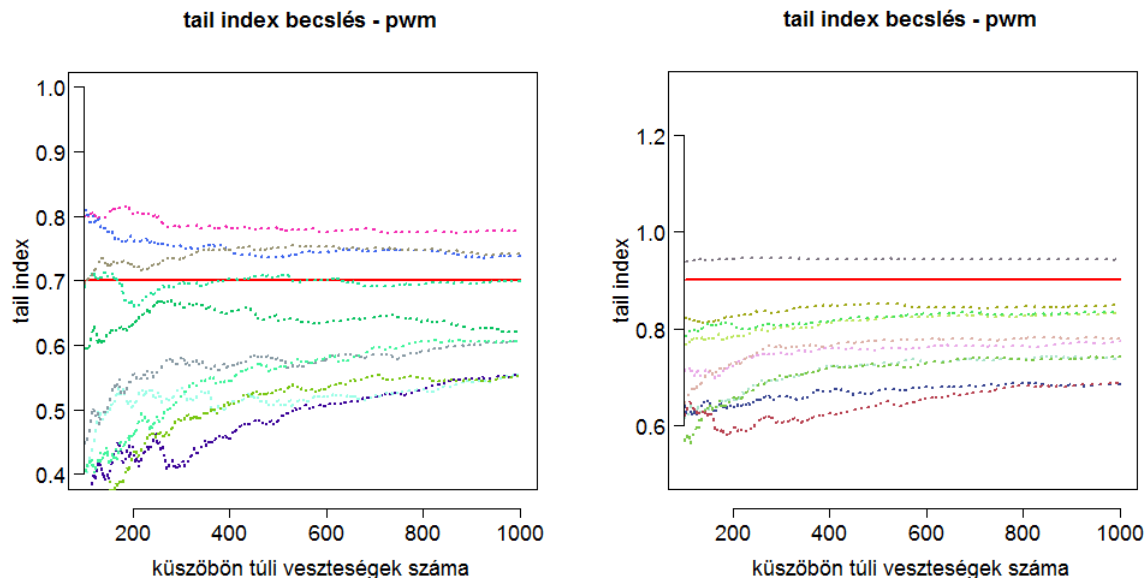


34. ábra: GPD eloszlás  $\xi$  paramétere a küszöb függvényében. Az ábra bal paneljei a többször előforduló értékek kiszűrése előtti becslést mutatják, míg a jobb oldali ábrákon a szűrést követő MLE becslés látható.

Látható, hogy a többszörösen előforduló értékek kiszűrése egy stabilabb MLE becslést eredményezett a fennmaradó mintán, a javulás leginkább ET7-ben szembetűnő. Mielőtt azonban következtetések vonnánk le a fenti becslés eredményéből, felvetjük a kérdést, hogy mennyit is ér a stabilitás? A PWM módszer esetében láthattuk, hogy egészen kis számú küszöbön túli veszteség esetén is stabil becslést kaptunk. Sőt nem kellett az MLE esetén jelen lévő, többszörösen előforduló értékek problémájával sem foglalkoznunk. A kérdés megválaszolásához a következő szimulációs kísérletet végeztük el. Adott paraméterek ( $\mu=1.000.000$ ;  $\beta=1.000.000$ ;  $\xi=[0,3 \ 0,5 \ 0,7 \ 0,9]$ ) paraméterek mellett 1000 elemű véletlen mintát generáltunk háromparaméterű GPD eloszlásból, és visszabecsültük a kétparaméteres GPD eloszlás scale paraméterét. Megjegyezzük, hogy a minta generálásakor a  $\mu$  (location parameter) biztosította, hogy a generált minta ne legyen  $\mu$ -nél kisebb, míg a visszabecslés során azt feltételeztük, hogy az adatok egy  $u = \mu$  küszöb fölött követnek GPD eloszlást, így már csak a kétparaméteres GPD eloszlás paramétereit kellett becslni. Fontos tudni, hogy a két GPD eloszlás közötti átváltás nincs hatással a  $\xi$  (scale) paraméterre (Moscadelli [2004]). A PWM szimulációs eredményeket a 35. ábra illusztrálja.



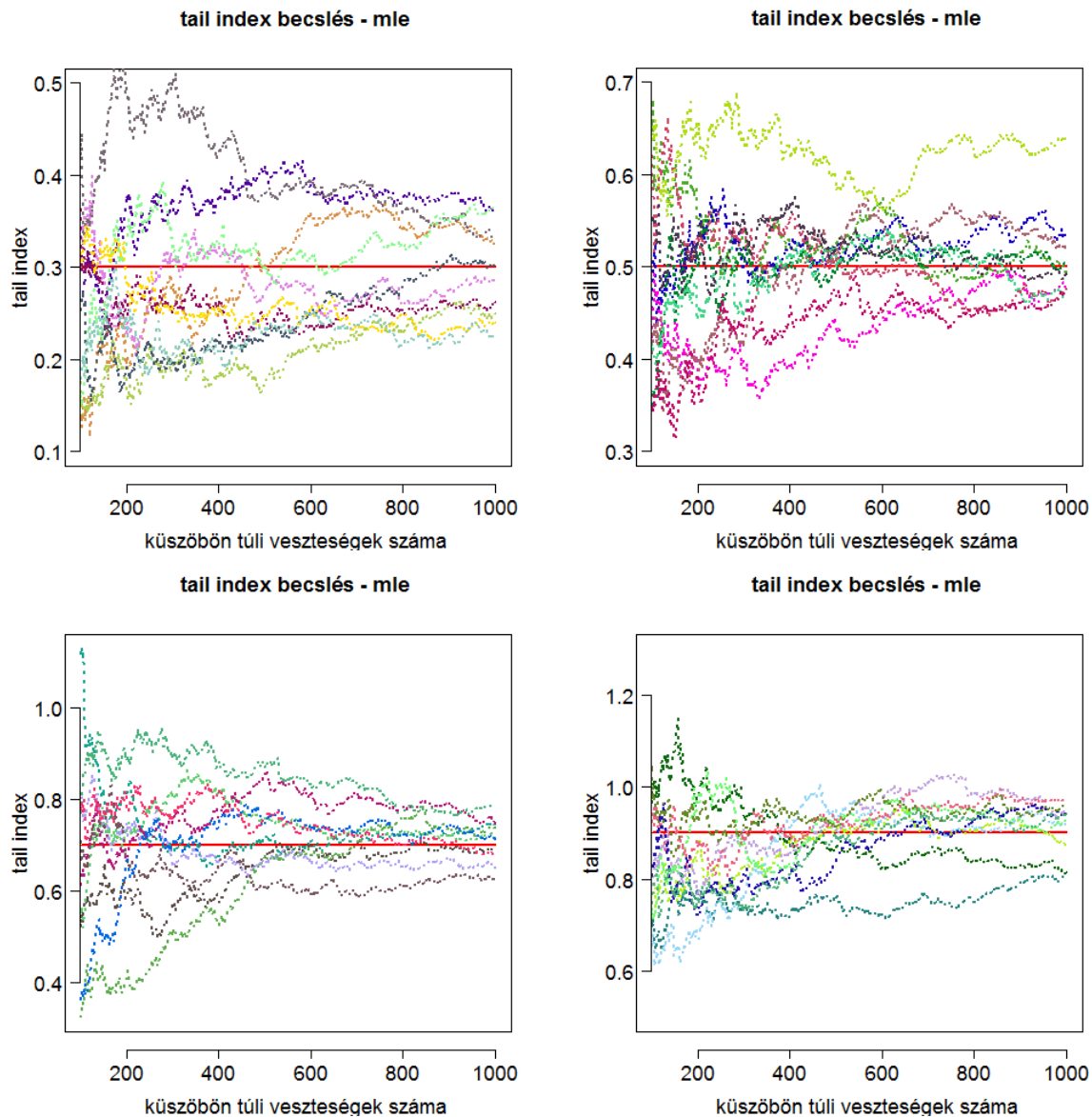




35. ábra: GPD eloszlásból generált mintából PWM módszerrel visszabecsült  $\xi$  paraméter a küszöb túli minta nagyságának a függvényében.<sup>36</sup>

Jól látható, hogy a PWM becslés stabilitása nem párosul a becslés jóságával. Ha a becslés rossz, azaz a becsült paraméter távol esik az igazi paramétertől, akkor az gyakran perzisztensen ott is maradt, a küszöbön túli veszteségek nagyobb száma egy adott szinten túl (top 300-400 mintaelemet meghaladóan) tipikusan nem, vagy nem lényegesen javítja a becslést a jelen szimulációs környezetben. Ennek függvényében még az is erősen megkérdőjelezhető, hogy elegendő-e a szakirodalom alapján javasolt minimális 100 elem a GDP paramétereinek a becsléséhez működési kockázatkezelési környezetben? Ugyanezt a kísérletet elvégeztük a MLE becslés keretében is, eredményét a 36. ábra illusztrálja.

<sup>36</sup> A piros folytonos vonal jelöli az igazi paramétert. Az egyes ábrák 10-10 szimulációs eredményt tartalmaznak. A paraméter becslését legalább 100 elemen végeztük el, tehát ennyi volt minimálisan a küszöbön túli veszteségek száma.



36. ábra: GPD eloszlásból generált mintából MLE módszerrel visszabecsült  $\xi$  paraméter a küszöbön túli minta nagyságának a függvényében.<sup>37</sup>

Az MLE becslés kapcsán is fellelhető az a jelenség, hogy a becslés a küszöbön túli veszteségek számától függetlenül rossz marad (konvergencia lelassul), ugyanakkor jól látszódik a módszer küszöbre való érzékenysége. Ezt leginkább az mutatja, hogy a PWM módszer esetében az egyes szimulációkból meghatározott paraméter útvonalak, főleg magasabb  $\xi$  mellett gyakorlatilag nem is keresztezik egymást (mintha „párhuzamosan” haladnának), addig az MLE becslés esetében ezt nem állíthatjuk.

<sup>37</sup> A piros folytonos vonal jelöli az igazi paramétert. Az egyes ábrák 10-10 szimulációs eredményt tartalmaznak. A paraméter becslését legalább 100 elemre végeztük el, tehát ennyi volt minimálisan a küszöbön túli veszteségek száma.

Az eddigiekből levonhatjuk azt a következtetést, hogy a tail index stabilitása nem feltétlenül jelenti azt, hogy megtaláltuk az igazi paramétert a tőkeszámításhoz, sőt éppen ellenkezőleg a  $\xi$  paraméter gyakran „ragad be” egy olyan állapotban, ami távol esik az igazi paramétertől. Ennek következtében erősen megkérdőjelezhetővé válnának a HUNOR adatbázis adatain becsült GPD jellemzők, így inkább arra helyezük a hangsúlyt a HUNOR paraméterek megállapítása helyett, hogy megismerjük a GPD eloszlás használatában rejlő instabilitást. Ugyanakkor még egy pillanatra visszatérve a HUNOR kategóriák tail index MLE becslésére, számos esetben tapasztaltuk, hogy a tail index helyenként, vagy tipikusan 1-nél nagyobb értéket vesz fel (ET2, ET3, ET4, ET6, ET7), legalábbis MLE becslés esetén. A dolgozat további részeiben még kitérünk ezen jelenség magyarázatára. Ha ez így van, akkor ez azt jelenti, hogy a GPD eloszlás egyik momentuma sem véges, várható értéke sem létezik. Mindez megkérdőjelezheti a GPD alkalmazhatóságát, hiszen éles ellentmondásban áll a probléma közgazdasági természetével. Azzal nem kívánunk foglalkozni, hogy a gyakorlatban milyen okai lehetnek annak, hogy a tail index becslése 1-nél nagyobb, ennek kapcsán az olvasó figyelmébe ajánljuk Neslehová et. al. [2006] cikkét.

Esetünkben az a kérdés, hogy a HUNOR keretei között (adott mintaelemszám) mellett, még ha minden modellfeltevés teljesülne is (pl. stacionaritás) meg lehetne-e érdemben becsülni a bank működési kockázati tőkéjét GPD módszertant alkalmazva? Bízhatnánk-e egy csak HUNOR adatokat felhasználó Bank LDA modelljében, illetve ha igen, akkor mennyire? Ezentúl kérdés, hogy az egyes modell bemeneti paraméterek milyen irányba befolyásolják (javítják, rontják) a becslés eredményét?

#### *8.6. GPD szimulációs kísérlet*

A fentiek eldöntésére újabb szimulációs kísérletet végzünk el, azzal céllal, hogy a modell keretein belül megismerjük az egyes paraméterek eloszlást, valamint hatását a működési kockázati tőkére. A kísérletben  $n=[200, 1000]$  elemű küszöbön túli külső adatbázisbeli mintákat generálunk, ami a HUNOR tekintetében nagyjából reálisnak nevezhető. A HUNOR statisztikákból láthattuk, hogy a rögzített veszteségek száma több esetben is csak pár száz, és itt még csak nem is vettük figyelembe azt a tényt, hogy az eloszlás széle csak egy adott küszöbön túl modellezhető GPD eloszlással, tehát a rendelkezésre álló minta reálisan még tovább csökken. Bizonyos esetekben 1000-et meghaladó mintaelemszámról beszélhetünk, de mivel a GPD csak limeszben teljesül, így ott is a mintanagyság csökkenésével kell

- az intézmény éves veszteségeinek az átlagos száma (N);
- külső adatbázis mintaelemszáma az adott veszteségkategóriában;
- az intézmény küszöbön túli (ahol a GPD már feltételezhető) veszteség számának az aránya a teljes mintán belül (H);
- becslési módszer (MLE, PWM) hatása a paraméterekre és a tőkére.

A szimulációs kísérlet kapcsán az SLA keretrendszert alkalmazzuk, mivel LDA környezetben minden egyes VaR értéket csak egy újabb időigényes szimulációval, vagy Fourier transzformációval<sup>38</sup> tudnánk előállítani. A kísérletben az előre definiált paraméterekkel, jellemzőkkel GPD eloszlású mintát generáltunk, majd MLE és PWM módszerekkel visszabecsültük a GPD paramétereit, és azok felhasználásával meghatároztuk a VaR értéket az Singe Loss Approximation segítségével.

A kapcsolódó ábrák mindegyike esetében a zöld szín az MLE, míg a piros szín a PWM becslésekhez tartozó eredményt jelenti. A paraméterek, vagy jellemzők igazi értékét minden esetben függőleges fekete szaggatott vonal jelzi. A zöld és piros függőleges szaggatott vonalak az igazi paraméter, vagy jellemző MLE és PWM becslését jelenti rendre. Az ábrák tartalmaznak még zöld és piros függőleges pontozott vonalakat, melyek 90 százalékos biztonsággal tartalmazzák az adott paramétert, vagy jellemzőt az MLE és a PWM becslés

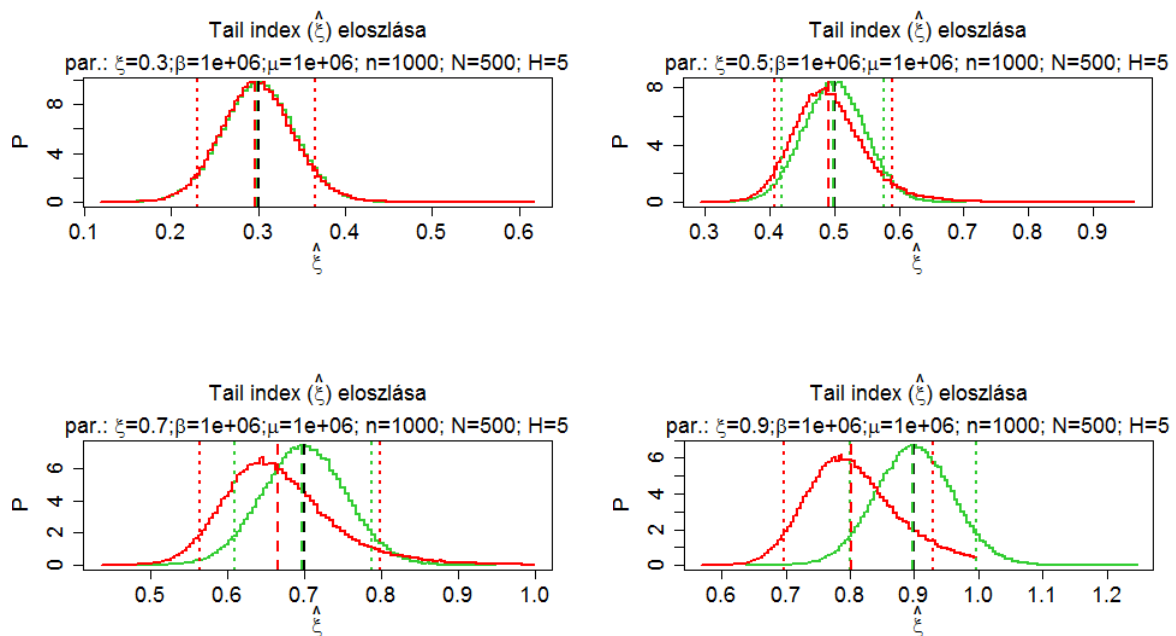
---

<sup>38</sup> Egyéb módszereket lásd korábban.

vonatkozásában. Az ábrákon túl számos további információt tároltunk el a paraméterekről, jellemzőkről, ilyen pl. azok szórása, ferdesége, terjedelme stb... Ezeket általánosságban nem tesszük közzé a dolgozatban, a lényeges információkat azonban közöljük.

### 8.6.1. A tail index eloszlásáról

A tail index kivételével rögzítettük a paramétereket, és a 37. ábra mutatja annak eloszlását.



37. ábra: Tail index eloszlása.

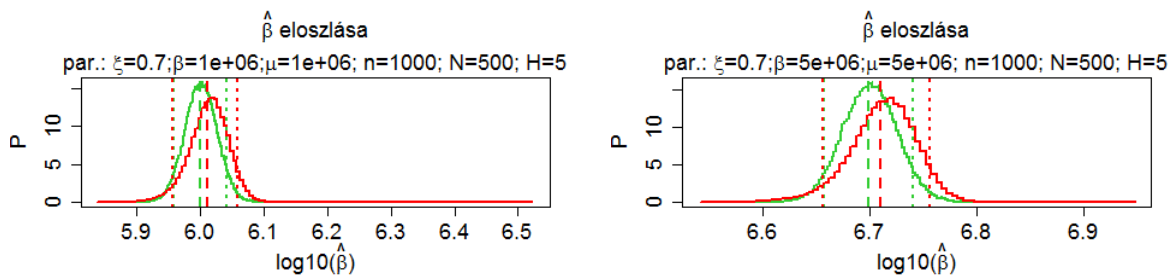
A szimulációs eredményekből a következőket olvastuk ki (illetve ismert):

- Az MLE becslés (aszimptotikusan) torzítatlanságát tükrözi vissza a szimuláció, míg a PWM becslőfüggvény várhatóan alulbecsli az igazi paramétert. A torzítás mértéke annál nagyobb (PWM esetén) minél közelebb van az igazi paraméter 1-hez.
- Az MLE becslés szimmetrikus, míg a PWM jobbra ferde.
- A PWM becslés felülről korlátos, míg az MLE nem! Ez összhangban van azzal, amit a HUNOR adatok esetében is láttunk, hogy a tail index mindig 1 alatt volt PWM becslés esetén.

- Minél közelebb van  $\xi$  értéke az 1 –hez annál inkább különbözik az MLE és a PWM becslés.
- MLE becslés esetén nagyjából 5% annak a valószínűsége, hogy a becslés során egy nem véges várható értékű eloszláshoz jutunk ( $\xi=0,9$  esetén).

### 8.6.2. A scale paraméter eloszlásáról

A scale paraméter vizsgálatának eredményét 38. ábra prezentálja.



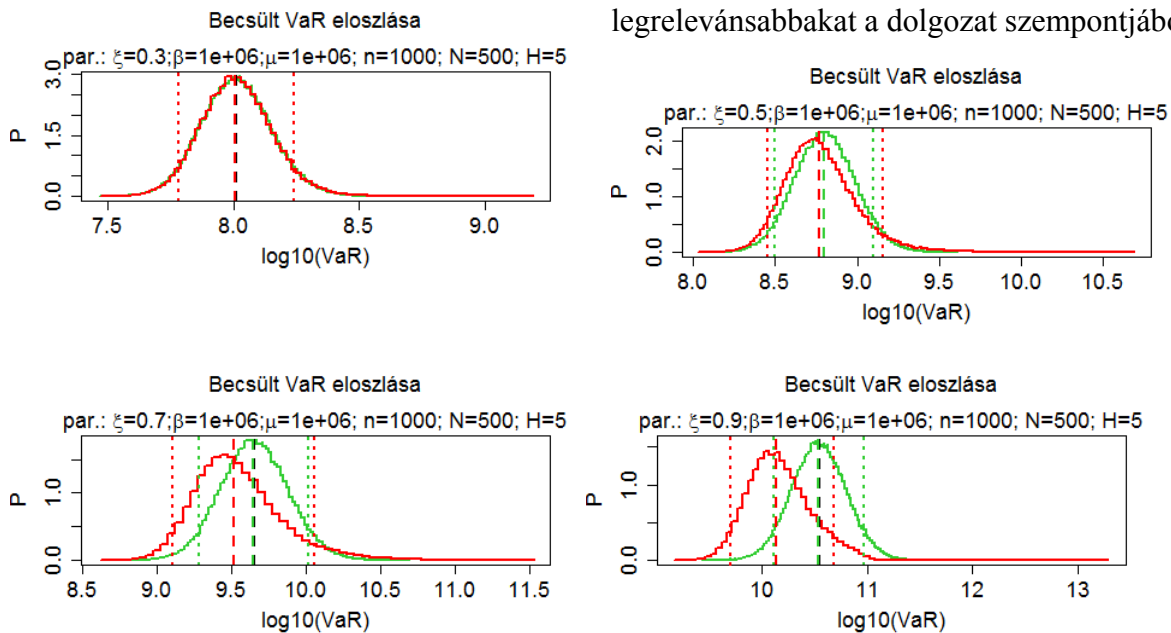
38. ábra: Scale paraméter eloszlása log10 skálán.

Megállapítások a szimuláció alapján:

- A scale paramétert a PWM becslőfüggvény felülbecsli, míg az MLE (aszimptotikusan) torzítatlan becslést ad.
- A 90%-os konfidencia sáv a PWM esetében szélesebb.
- Mérésckelt a különbség az egyes becslőfüggvényekhez tartozó eloszlások között, ezt erősíti meg a paraméter becslési hibájának a tőkére gyakorolt hatása is (lásd később).
- A nagyobb  $\beta$  nagyobb szórással is jár, de a relatív szórások már megegyeznek.

### 8.6.3. A tőke eloszlásáról

Számunkra a leginkább izgalmas kérdés, hogy a becslési hiba következtében hogyan alakul a tőke? A korábban vázolt szempontok mindegyike befolyásolja a tőkét, így itt sem teszünk közzé minden eredményt, csak a legrelevánsabbakat a dolgozat szempontjából.



39. ábra: VaR eloszlása log10 skálán.

A VaR szimulációja során az alábbi lényeges vonások kerültek felszínre:

- A PWM becslőfüggvény a tőkét várhatóan alulbecsli, míg az MLE becslés azt aszimptotikusan torzítatlan módon becsl.
- A 90%-os biztonsággal becsült tőke alsó és felső értéke között megközelítőleg 1 nagyságrendnyi eltérés van (az adatok log10 skálán vannak feltüntetve), még mérsékelt tail index mellett is!
- A 98%-os biztonsággal becsült tőke minimuma és maximuma között 30-szoros eltérés is előfordult (lásd 29. táblázat: PWM,  $\xi=0,7$  esetén)! A maximum és minimum hányadosa minden esetben kisebb az MLE becslés esetén, mint a PWM módszer használatával. Szembetűnő azonban az a jelenség, hogy  $\xi=0,9$  esetén a PWM becsléshez tartozó VaR eloszlásának felső (0,99) kvantilise kisebb, mint az MLE becsléshez tartozó VaR ugyanazon kvantilise. Ennek oka a  $\xi$  eloszlásának vizsgálatakor megfigyelt jelenség lehet, hogy a PWM becslés során  $\xi$  maximális értéke 1 volt, tehát a PWM becslés sosem eredményezett olyan  $\xi$  értéket, ami mellett a GPD eloszlás várható értéke ne létezne.

M esetében érdemes megfigyelni, hogy a szórás a várható érték többszörösére nő  $\xi=0,9$  esetén!

- A PWM-hez tartozó VaR eloszlása jobbra ferde. A ferdeség annál inkább jellemzőbb vonássá válik minél közelebb van  $\xi$  az 1-hez.

Tail index	Várható érték	Szórás	Ferdeség	1% -os kvantilis	99% -os kvantilis
0,300	107.152.810	35.740.310	1,163	49.413.672	220.351.285
0,500	682.343.972	308.406.733	1,563	237.403.592	1.709.173.438
0,700	5.068.958.461	2.841.043.346	1,970	1.364.184.519	14.975.533.496
0,900	41.226.445.660	27.636.172.735	3,074	8.869.118.172	140.913.485.692

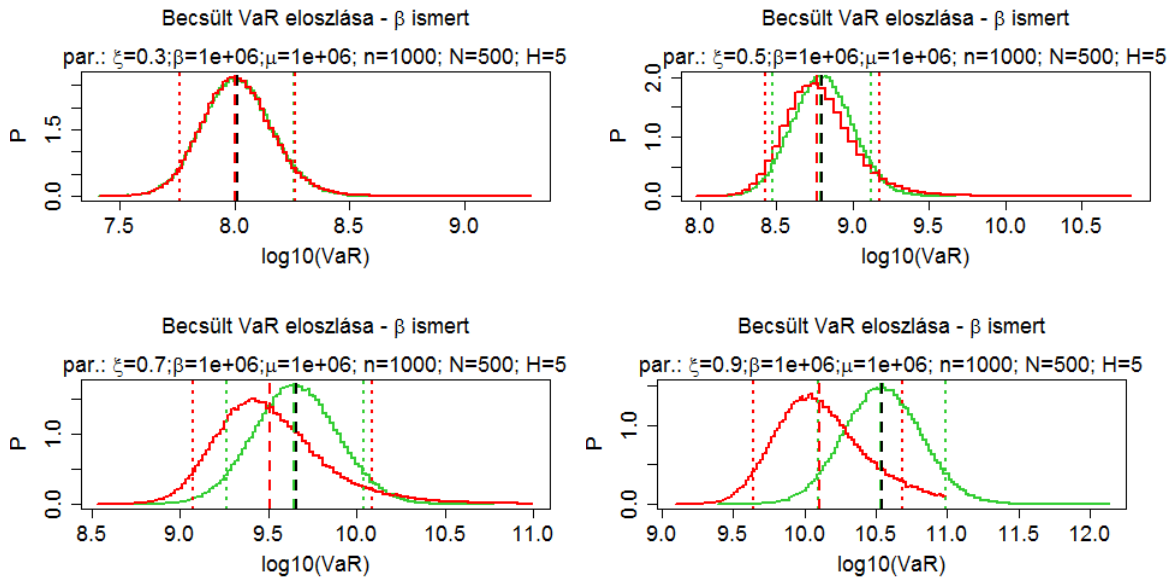
28. táblázat: VaR jellemzők MLE becslőfüggvénnyel. Az egyéb, táblázatban fel nem tüntetett paraméterek megegyeznek a 39. ábra paramétereivel.

Tail index	Várható érték	Szórás	Ferdeség	1% -os kvantilis	99% -os kvantilis
0,300	106.806.920	39.690.590	3,955	49.495.505	232.497.014
0,500	697.585.672	822.136.970	20,716	216.612.281	2.826.418.116
0,700	4.387.413.672	5.220.227.681	79,650	930.914.271	27.425.964.614
0,900	18.138.190.165	72.316.704.424	225,115	3.625.493.635	76.118.872.478

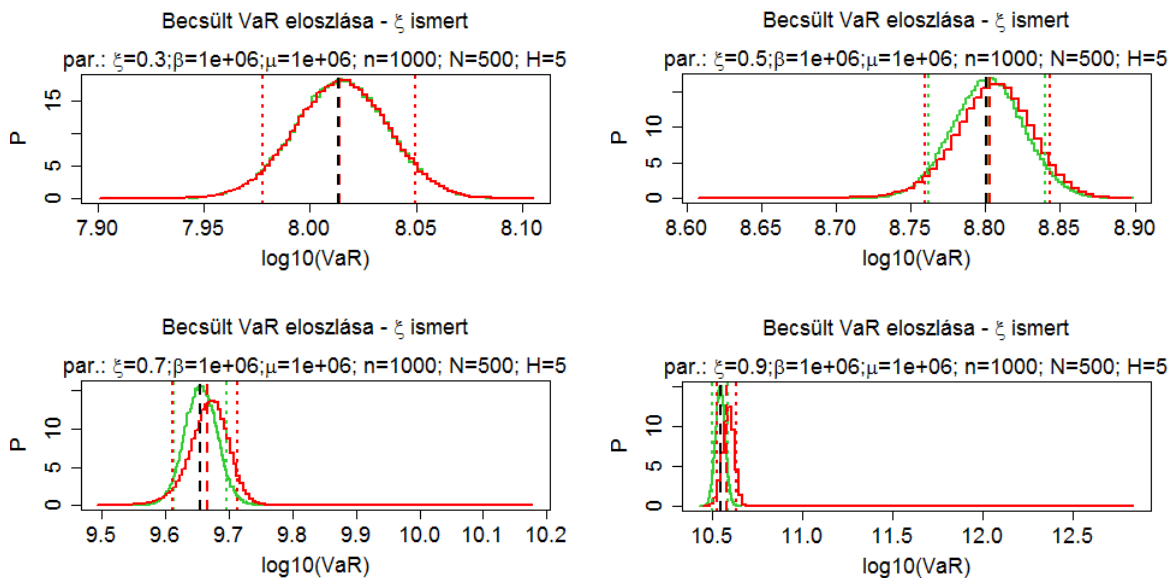
29. táblázat: VaR jellemzők PWM becslőfüggvénnyel. Az egyéb, táblázatban fel nem tüntetett paraméterek megegyeznek a 39. ábra paramétereivel.

A 39. ábra VaR eloszlása magában foglalja mind a  $\xi$  mind a  $\beta$  paraméter becslésének hibáját. Ha párhuzamosan rápillantunk a 37. ábra és 39. ábra eloszlásaira, az a benyomásunk támad, hogy a két ábra eloszlásai rendre nagyon hasonlók, és elsődlegesen  $\xi$  becslési hibája határozza meg a tőke eloszlását. Ennek alátámasztására bemutatjuk a VaR eloszlását előbb igazi  $\xi$ , majd az igazi  $\beta$  paraméterek mellett.





40. ábra: VaR eloszlása log10 skálán,  $\beta$  igazi értéke mellett.

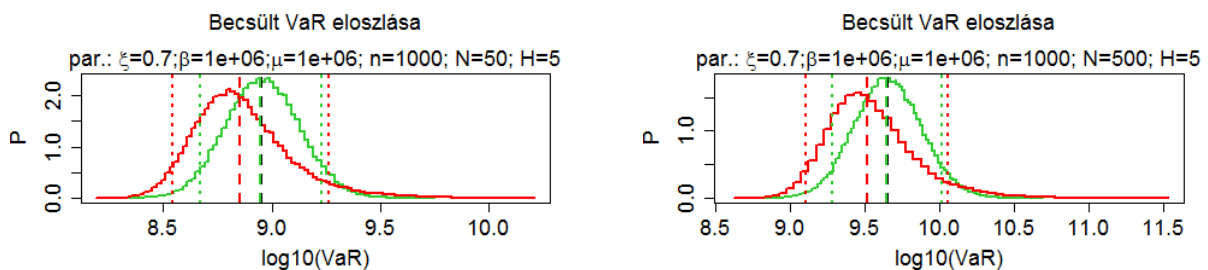


41. ábra: VaR eloszlása log10 skálán,  $\xi$  igazi értéke mellett.

A 40. ábra és 41. ábra összehasonlítása alapján jól látszódik, hogy  $\xi$  becslési hibája lényegesen nagyobb hatással van a tőkére, mint  $\beta$  becslési hibája. Ez megerősít minket abban a hitünkben, hogy a Pareto eloszlás illesztése során  $\xi$  becslésére kell nagyobb hangsúlyt fektetni, illetve a VaR becslésben elkövetett hiba elsődlegesen a  $\xi$  becslésében elkövetett hibára vezethető vissza.

A következő bemeneti paraméter, melynek a tőkére való hatását vizsgáltuk az intézmény éves gyakoriság értéke. Mivel a kísérletek a Single Loss Approximation keretei között zajlottak,

így csupán az éves gyakoriság várható értékét használtuk fel, és képezte a szimulációk inputját. Két esetet vizsgáltunk, az  $N=[50\ 500]$  értékekkel. Az  $N=500$  nagynak tűnhet még az  $n=1000$ -hez képest is, amennyiben úgy gondolunk rá, hogy az intézmény a külső adatokkal egy az egyben bővíti adatbázisát. Ez esetben ha már 5 hasonló méretű bank van a piacon (amint az a HUNOR esetében is fennáll), akkor legalább  $n=2500$  elemmel kellett volna kísérletet elvégezni, még ha csak 1 évnyi minta áll is a bank rendelkezésére. A felépített kísérlet azonban a  $H$  értéken keresztül szabályozza a GPD-ként figyelembe vehető mintaelemszámot, ami  $H=50$  esetén 2%, míg  $H=5$  esetén 20%. Előbbi esetén az 5 hasonló méretű intézmény évente, egyenként 10, míg a második esetben 100 GPD mintaelemet szolgáltat a közös adatbázisba. Ennek függvényében az  $N=500$  már nem aránytalanul nagy az  $n=1000$ -hez képest. Mint azt már korábban említettük nem nyilvánvaló, hogy az extrapolációs probléma miatt a több, vagy kevesebb belső adattal rendelkező intézmények járnak-e jobban? Erre lognormális eloszlású veszteségek esetén ugyan adhatunk egyfajta választ, ugyanakkor az intézmény itt egy vastag szélű eloszlással szembesül a korábbi semi-heavy tail helyett, illetve a belső adatok számossága nincs kapcsolatban a külső adatbázis méretével. Ennek következményeként a kérdést a jelen kísérlet keretei között ismételtelen felvetettük.



42. ábra: VaR eloszlása log10 skálán, az éves gyakoriság ( $N$ ) különböző értékei mellett.

Amint azt az ábrák mutatják az MLE és PWM VaR eloszlások egymáshoz viszonyított helyzete igen hasonló különböző gyakoriságok mellett. Az összehasonlító táblázatokból kiderül, hogy a nagyobb gyakorisággal rendelkező intézmény nem feltétlenül van jobb helyzetben. Az éves gyakoriság növelésével a VaR relatív szórása (szórás / VaR hányados) nőtt mind az MLE, mind a PWM esetében,  $\xi$  minden értékére! Ugyanakkor a relatív szórás mértéke az MLE becslés esetében a kisebb, nagy  $\xi$  esetén, csaknem egy nagyságrenddel. A várható értékkel normált tartományok, melybe a VaR 98%-os valószínűséggel esik, mind az MLE mind a PWM becslés esetén nőttek! Felhívjuk a figyelmet, hogy az extrapolációs problémának most egyértelműen a negatív hatásával szembesül egy sok belső adattal

rendelkező intézmény egy kevés adattal rendelkezőhöz képest. A nagyobb belső mintaelemszám miatt távolabbra kell extrapolálnia az eloszlás szélére a VaR megállapításához, mint egy kevesebb belső adattal rendelkező intézménynek, miközben mindkettőjüknek ugyanakkora minta (külső adatbázis) áll rendelkezésre az eloszlás szélén! Ez pedig rontja a becslés jóságát.

Tail index	Éves gyakoriság	Várható érték	Szórás	Q <sub>01</sub>	Q <sub>99</sub>	R <sub>01</sub>	R <sub>99</sub>
0,300	50	51.514.350	12.418.379	29.584.524	88.095.944	0,574	1,710
0,300	500	107.152.810	35.740.310	49.413.672	220.351.285	0,461	2,056
0,500	50	207.163.988	67.863.082	96.596.593	417.271.056	0,466	2,014
0,500	500	682.343.972	308.406.733	237.403.592	1.709.173.438	0,348	2,505
0,700	50	962.220.506	398.276.581	365.034.651	2.258.012.617	0,379	2,347
0,700	500	5.068.958.461	2.841.043.346	1.364.184.519	14.975.533.496	0,269	2,954
0,900	50	4.847.168.669	2.377.139.947	1.522.753.546	12.885.649.931	0,314	2,658
0,900	500	41.226.445.660	27.636.172.735	8.869.118.172	140.913.485.692	0,215	3,418

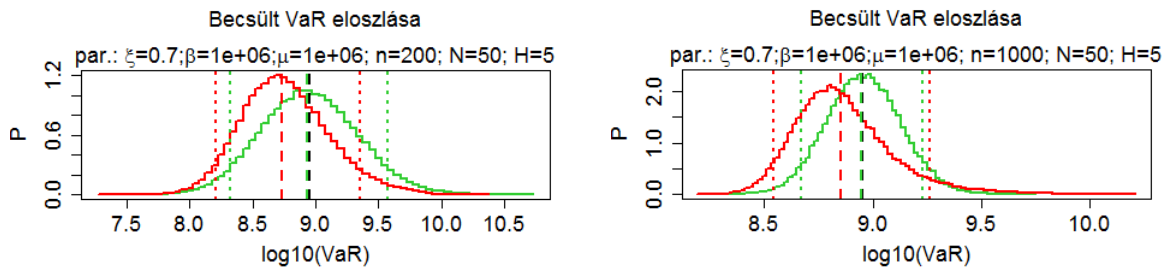
30. táblázat: VaR jellemzők MLE becslőfüggvénnyel. Az egyéb, táblázatban fel nem tüntetett paraméterek megegyeznek a 39. ábra paramétereivel.

Tail index	Éves gyakoriság	Várható érték	Szórás	Q <sub>01</sub>	Q <sub>99</sub>	R <sub>01</sub>	R <sub>99</sub>
0,300	50	51.335.693	13.404.257	29.633.966	92.032.112	0,577	1,793
0,300	500	106.806.920	39.690.590	49.495.505	232.497.014	0,463	2,177
0,500	50	206.454.416	131.300.400	89.907.780	615.045.253	0,435	2,979
0,500	500	697.585.672	822.136.970	216.612.281	2.826.418.116	0,311	4,052
0,700	50	832.197.012	616.178.472	276.862.252	3.538.222.286	0,333	4,252
0,700	500	4.387.413.672	5.220.227.681	930.914.271	27.425.964.614	0,212	6,251
0,900	50	2.539.641.308	1.793.744.629	785.956.266	8.067.591.458	0,309	3,177
0,900	500	18.138.190.165	72.316.704.424	3.625.493.635	76.118.872.478	0,200	4,197

31. táblázat: VaR jellemzők PWM becslőfüggvénnyel. Az egyéb, táblázatban fel nem tüntetett paraméterek megegyeznek a 39. ábra paramétereivel.

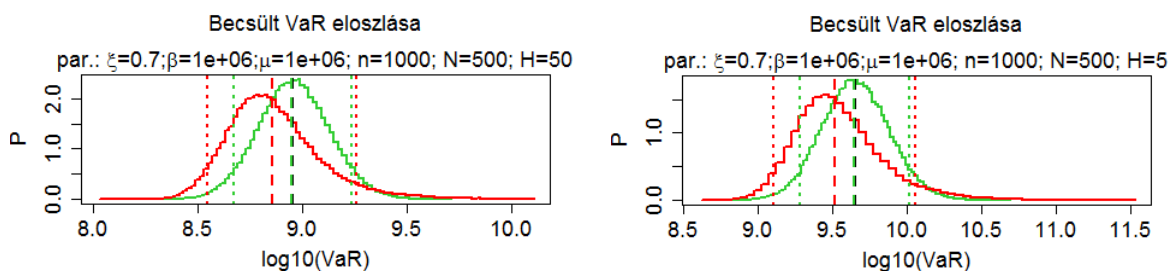
Vizsgálatunk következő célpontja a külső adatbázis mérete. Az előbbieken már rávilágítottunk arra, hogy a külső adatbázis használata kedvezőbb egy kis intézmény számára, mint egy nagyobb számára. Ugyanakkor a külső adatbázis veszteségeinek a gyarapodása mind

a kevés mind a sok veszteséggel rendelkezőket egyre kedvezőbb helyzetbe hozza (feltéve, hogy a veszteség generáló folyamat időben nem változik). A 43. ábra jól mutatja, hogy nagyobb külső minta ( $n$ ) esetén a VaR eloszlása koncentráltabb, az kisebb tartományban szóródik.



43. ábra: VaR eloszlása log10 skálán, a külső adatbázis különböző mérete ( $n=[200\ 1000]$ ) mellett.

Az utolsó, amit a kísérletek során vizsgáltunk, a  $H$  jellemző. Ez jelenti az összes veszteség számának és a küszöbön túli veszteségek számának az arányát. Ezt az adott pénzügyi intézmény a belső adatai alapján határozza meg, és ezáltal egyben azt is determinálja, hogy mekkora súlyt kapnak a külső adatok. Ha értéke nagy, akkor csak az eloszlás legszélét feltételezi a bank GPD eloszlásnak, míg kisebb értékei esetén a bank már az eloszlás törzséhez közeli részeket is Pareto eloszlásúnak feltételezi. Az eredmények megismerése előtt azt gyaníthatjuk, hogy  $H$  nagyobb értéke (ami egyúttal kevesebb extrém értéket is jelent a belső mintában) a kedvezőbb a bank számára. Ennek oka, hogy ez esetben az egész GPD eloszlás kisebb súlyt kap a teljes eloszlásban, így már kevésbé kell a GPD eloszlásban extrapolálni, hogy az SLA szerinti valószínűséghez tartozó kvantilist meghatározzuk.



44. ábra: VaR eloszlása log10 skálán, a  $H$  jellemző különböző értékei mellett.

A részletező táblázat nélkül is jól látható, hogy  $H$  kisebb értéke nagyobb tőkét eredményez a pénzügyi intézmény számára.

Megjegyezzük, hogy ez a példa annyiban nem életszerű, hogy az  $u$  küszöb változatlan a szimulációkban. Gyakorlatban, ha egyre több extrém veszteséget vonunk a GPD vonzástartományába, az azt jelenti, hogy  $u$  küszöböt egyre inkább csökkentem, hogy a veszteségek a GPD eloszláshoz tartozhassanak. A kísérletben azonban az  $u$  küszöb változatlan maradt, így pusztán azt feltételeztük, hogy az egyik intézménynek kevesebb, míg a másiknak több eseménye van ugyanazon  $u$  küszöb felett. Mivel a tőkét elsődlegesen az extrém veszteségek határozzák meg, így nem meglepő, hogy a több extrém veszteséggel rendelkező intézmény képez nagyobb tőkét.

### *9. Konklúziók a működési kockázati tőkekövetelmény instabilitásáról*

A dolgozat során arra törekedtünk, hogy mélyebben megértsük annak a működési kockázatkezelési keretrendszernek a kihívásait, melyet a Basel II szabályozás hozott életre, és ami lehetővé tette a belső modellek alkalmazását a működési kockázatkezelésben.

A dolgozat bevezető fejezeteiben összefoglaltuk az AMA módszertannal kapcsolatos problémák lényeges vonásait. Ennek kapcsán kitértünk számos olyan problémára, ami nehézséget jelent egy AMA modell bevezetésekor és alkalmazásakor. Felvettük a Loss Distribution Approach és Single Loss Approximation módszertanának elméleti alapjait és részletesen tárgyaltuk azt, hogy a pénzüintézeteknek milyen feladatokat kell elvégezni ezek alkalmazás során.

Számos szerző felhívta a figyelmet az AMA modellekkel kapcsolatos problémákra, és köztük elsődlegesen a tőkekövetelmény instabilitására, ugyanakkor az ezt magyarázó elsődleges okot leginkább a Single Loss Approximation elméletén keresztül lehet megérteni. Eszerint az intézmények sosem rendelkeznek elegendő információval a tőke becsléséhez, 5 éves idősor esetén 200-szoros extrapoláció szükséges a tőke meghatározásához. Ez azonban azt sugallja, hogy teljesen mindegy, hogy kis (kevés belső adattal rendelkező) vagy nagy (sok belső adattal rendelkező) intézménnyel van-e dolgunk, a tőke becslése egyaránt bizonytalan. Annak eldöntésére, hogy az ilyen értelemben vett intézményméret számít-e, egy modellt építettünk, ami kezelhetőbbé tette a problémát. A modellépítés elsődleges hozadéka, hogy analitikusan meghatározhatóvá válik a tőke eloszlása, annak minden jellemzőjével együtt. Ez lényegesen leegyszerűsíti azt a hagyományos utat, ami nagyméretű szimulációkon, diszkretizálási feladatokon, vagy komplex numerikus eljárásokon (FFT) keresztül teszi

megismerhetővé a tőkebecslésben rejlő bizonytalanságot. Az adott modellfeltételek mellett az alábbi lényeges következtetéseket vontuk le:

- A tőke eloszlása nem szimmetrikus, hanem jobbra ferde.
- Az egyedi lognormális veszteségeloszlású modellből becsült parametrikus VaR eloszlásának várható értéke és szórása egyaránt függ a mintaelemszámtól ( $n$ ). Várható értékben az intézmények túlbecslik a tőkeszükségletüket. Kis gyakoriság mellett a túlbecslés várhatóan nagyobb.
- A becslés hibája (VaR szórása) 5 éves idősről, kevésbé heavy tail ( $\sigma$  kicsi) esetben monoton csökkenő a gyakoriság függvényében, míg növekvő  $\sigma$  mellett a kezdeti csökkenés megtorpan és  $\lambda$  magasabb értékei mellett újra emelkedni kezd. Vastag szélű eloszlás esetén a szórás monoton növekvő a gyakoriság függvényében. Ugyanakkor ez a jelenség már nem áll fenn a relatív szórás esetében, az csökkenő tendenciát mutat a gyakoriság függvényében, azaz a nagyobb intézmény méret (több adat) kisebb becslési hibával járul.

A fentiekén túl elemeztük a HUNOR adatbázisban rejlő lehetőségeket is. Itt elsődlegesen az általánosított Pareto eloszlásra koncentráltunk. A külső adatok használata csökkenti az extrapolációs problémát, így arra voltunk kíváncsiak, hogy a HUNOR adta lehetőségek (gyakoriság szám) mellett érdemben csökkenthető-e a tőke becsléséhez társuló bizonytalanság, és ha igen, akkor az milyen nehézségekkel jár? Leíró elemzésünk során megállapítottuk, hogy a HUNOR adatbázis veszteségei valóban vastag szélű eloszlásból származnak, illetve hogy bizonyos kategóriák alig tartalmaznak néhány tíz, vagy száz adatot. Elvégeztük az egyes kategóriák MLE és PWM becslését, és az alábbi észrevételeket tettük:

- A mintaelemszám függvényében elvégzett tail index MLE becslés fűrészfogszerű képet mutat, melynek oka az egyes kategóriák azonos értékei. Ezek figyelembe vétele mindenképp szükséges egy tőkemodell megalkotásakor.
- A PWM becslés sima és stabil becslést eredményezett.
- Az MLE becslés gyakran 1 fölötti tail indexet eredményezett, ami a veszteséggeneráló folyamat nem véges várható értékére utal.

Mivel a fenti pontok az első kivételével további vizsgálatot igényeltek, így mesterséges, szimulációs környezetben vizsgáltuk tovább azokat. Itt megvizsgáltuk a GPD eloszlás paramétereinek eloszlását és tőkére gyakorolt hatását, valamint néhány olyan jellemzőt, mint mintaelemszám, becslési módszer, melyek szintén kritikusak lehetnek a tőke meghatározásakor. A főbb eredmények a következők:

- A GPD paraméterek közül a tail index hatása a tőkére lényegesen nagyobb, mint a shape paraméteré.
- A szimuláció az MLE becslés aszimptotikus torzítatlanságát visszaigazolta a tail index becslése során, míg a PWM becselőfüggvény várhatóan alulbecsli az igazi paramétert. A torzítás mértéke annál nagyobb (PWM esetén) minál közelebb van az igazi paraméter 1-hez.
- Az MLE tail index becslés szimmetrikus, míg a PWM jobbra ferde.
- A PWM tail index becslés felülről korlátos, míg az MLE nem! Ez összhangban van azzal, amit a HUNOR adatok esetében is láttunk, hogy a tail index mindig 1 alatt volt PWM becslés esetén.
- Minél közelebb van  $\xi$  értéke az 1 –hez annál inkább különbözik az MLE és a PWM becslés.
- A 90%-os biztonsággal becsült tőke alsó és felső értéke között megközelítőleg 1 nagyságrendnyi az eltérés még mérsékelt tail index mellett is! A 98%-os biztonsággal becsült tőke minimuma és maximuma között 30-szoros eltérés is előfordult (miközben az intézmény 500 belső adattal és 1000 külső adattal rendelkezett az adott kategóriában)! A maximum és minimum hányadosa minden esetben kisebb az MLE becslés esetén, mint a PWM módszer használatával.
- A relatív szórás nő, ahogy  $\xi$ -vel 1-hez közelítünk. PWM esetében érdemes megfigyelni, hogy a szórás a várható érték többszörösére nő  $\xi=0,9$  esetén!
- A PWM-hez tartozó VaR eloszlása jobbra ferde. A ferdeség annál inkább jellemzőbb vonássá válik minél közelebb van  $\xi$  az 1-hez.
- A nagyobb gyakorisággal rendelkező intézmény nem feltétlenül van jobb helyzetben a kisebbekhez képest. Az éves gyakoriság növelésével a VaR relatív szórása (szórás /

VaR hányados) nőtt mind az MLE, mind a PWM esetében,  $\xi$  minden értékére! Ugyanakkor a relatív szórás mértéke az MLE becslés esetében a kisebb, nagy  $\xi$  esetén, csaknem egy nagyságrenddel. A várható értékkel normált tartományok, melybe a VaR 98%-os valószínűséggel esik, mind az MLE mind a PWM becslés esetén nőttek nagyobb gyakoriság mellett! Felhívjuk a figyelmet, hogy az extrapolációs problémának most egyértelműen a negatív hatásával szembesül egy sok belső adattal rendelkező intézmény egy kevés adattal rendelkezőhöz képest. A nagyobb belső mintaelemszám miatt távolabbra kell extrapolálnia az eloszlás szélére a VaR megállapításához, mint egy kevesebb belső adattal rendelkező intézménynek, miközben mindkettőjüknek ugyanakkora minta (külső adatbázis) áll rendelkezésre az eloszlás szélén! Ez pedig rontja a becslés jóságát.

Összességében láthatjuk, hogy AMA modell építéskor az intézménynek igen körültekintően kell eljárnia, hiszen egy olyan szabályozói környezetben kell a tőkét meghatároznia, ahol a tőkeérték messze a megfigyelhető és visszamérhető tartományon túl helyezkedik el. Láthattuk, hogy a nagy intézmények sem élveznek feltétlenül előnyt a tőke stabilitást illetően, hiszen míg a nagyobb belső adatbázissal ők tudják jobban csökkenteni a becslésben rejlő hibát, addig a külső adatbázis fix mérete a kisebb intézményeknek kedvez, ahol nem szükséges annyira távolra extrapolálni a tőke meghatározásához.

A konklúziókat Richard Smith gondolatával zárjuk, miszerint „There is always going to be an element of doubt, as one is extrapolating into areas one doesn't know about.”



## 10. Portfóliók zajérzékenysége

(Kutatási hipotézis)

Egy másik igen fontos kockázatkezelési probléma, melynek kapcsán számos instabil jelenség figyelhető meg a portfólióoptimalizáláshoz kötődik. Így a korábbi működési kockázati kontextustól elszakadva röviden bemutatjuk Markowitz portfólióoptimalizálási keretrendszerét, és alkalmazzuk azt napjaink kockázati mértékeire. Ennek kapcsán látni fogjuk, hogy az instabilitás egyrészt a portfóliósúlyok nagyfokú változékonyságában, másrészt a portfólió optimalizálhatóságában tükröződik vissza.

Habár a pénzügyek egyik központi kérdése a kockázat, mely fontos szerepet játszik a befektetési döntéseknél, eszközök allokációjánál, a kockázatmenedzsmentnél és a szabályozásoknál, alapvető fontossága ellenére nincs egyetemesen elfogadott mérték annak kvantitatív jellemzésére. A kockázati mértékeket a gyakorló pénzügyesek kockázatkezelési szoftverekben implementálják, különböző szabályozásokban kezelik, és elméleteket gyártanak a hüvelykujjszabályon, vagy ad hoc recepteken keresztül a kifinomult axiomatikus konstrukciókig. A kockázati mértékeket a gazdaságelmélet szélesebb kontextusába ágyazva szokás, (vagy legalábbis így kellene) kezelni. Ez leginkább a hasznosság függvényekkel való kapcsolatukat jelenti, ugyanakkor nem nyilvánvaló, hogy a kockázati mérték választása hogyan befolyásolja befektetőket és piacok természetét. Ezen dolgozat során gyakorlati szempontból nézünk meg néhány konkrét kockázati mértéket, melyek a gyakorlatban, szakirodalomban, és szabályozásokban elterjedtek, és egyetlen dologra fókuszálunk, mégpedig a becslésük során elkövetett hiba mértékére, azaz zajérzékenységükre<sup>39</sup>. Látni fogjuk, hogy minden kockázati mérték jelentős zajérzékenységet mutat, és a különböző kockázati mértékek ugyanazon zajra nagyon különbözőképpen reagálnak. Azt is megmutatjuk, hogy a kifinomultabb (újabb) kockázati mértékek érzékenysége magasabb, mint a konvencionális kockázati mértékeké.

A legtöbb gyakorló szakember a kockázatot egy adott számként, vagy néhány szám együtteseként kezeli. Ám ez egy leegyszerűsített nézőpont. A kockázatkezelési információkat a becslések eredménye jelenti, melyekhez piaci empirikus megfigyelésekből nyerjük az inputot. Mivel a vizsgált minták mindig végesek, mintáról mintára való fluktuációjuk elkerülhetetlen. Bármely pénzügyi eszköz vagy portfólió kockázati jellemzői ezért sosem jól definiálható értékek, hanem valószínűségi változók. Ideális esetben ezek eloszlása annyira

---

<sup>39</sup> A fogalom bevezetése a közös publikációk során Kondor Imre és Pafka Szilárd fizikusoktól származik.

csúcsos lenne, hogy egyetlen minta a teljes eloszlást jól reprezentálná. Tény azonban, hogy a tipikus banki portfóliók mérete, és az elérhető idősorok hosszúsága olyan, hogy ez általában sosem valósul meg, és a becslési hiba problémája pénzügyi szempontból soha nem hagyható figyelmen kívül. Az alapvető probléma az információhiány. Az az adatmennyiség, mely pontosan rekonstruálná az alapul szolgáló sztochasztikus folyamatot, jelentősen meghaladja a tipikus esetben rendelkezésre álló adatmennyiséget. Ilyen körülmények között sztochasztikus modellünk elkerülhetetlenül tartalmaz több-kevesebb mérési hibát, és a kockázat becslése zajos lesz.

Hangsúlyozzuk, hogy egy kockázati mértéket két célra használhatunk. Egyrészt mérhetjük vele egy adott portfólió kockázatát, másrészt döntéseket hozhatunk a mérték felhasználásával kapcsolatosan (portfólióoptimalizálás). A zaj hatása sokkal komolyabb, amikor adott feltételek mellett szeretnénk kiválasztani az optimális portfóliót, mint akkor, amikor csupán egy meglévő portfólió kockázatát becsüljük. Ugyanazon kockázati mérték kielégítőnek bizonyulhat diagnosztikus eszközként, de gyengén teljesíthet döntéstámogató eszközként. Ezen dolgozat a döntéstámogató szerepre fókuszál. Nem a portfólió kockázatát becsüljük, hanem a kockázati mértékek portfólióválasztás során tanúsított zajérzékenységét.

Ez a problematika természetesen nem új keletű. Felfedezése Markowitz racionális portfólióválasztási elméletének legelejére tehető (Markowitz [1952], Markowitz [1959]), és az évtizedek alatt rengeteg cikk foglalkozott vele különböző szemszögből. A szakirodalom nagy része olyan portfóliókkal foglalkozik, amelyek Gauss statisztikát követnek, és a Gauss portfóliókra vonatkozó természetes kockázati mértéknek, a varianciának a zajérzékenységét vizsgálják (ld. Frankfurter et. al. [1971]; Dickinson [1974]; Jobson et. al. [1980]; Elton et. al. [1973]; Eun et. al. [1984]; Chan et. al. [1999]). A zajhatás tanulmányozása mellett számos zajcsökkentő rendszert is kifejlesztettek, hogy a portfólióválasztás során a becslési hiba kisebb - nagyobb részét kiiktassák. Ezen szűrőeljárások többek között egy, vagy többfaktorú modelleket (Elton et. al. [1995]), Bayesi becsléseket, vagy pl. a véletlen mátrix elmélet eszközeit tartalmazzák.

Annak ellenére, hogy az irodalomban különböző kockázati mértékeken alapuló portfólióoptimalizálási módszerek kerültek bemutatásra (Konno et. al. [1991]; Young [1998]; Rockafellar et. al. [2000]; Acerbi [2004]), és ezek közül néhányat széleskörűen alkalmaznak a gyakorlatban (Dembo et. al. [2000]; Algorithmics [2002]), a variancián kívüli más kockázati mértékek zajhatásaival foglalkozó szakirodalom sokkal szűkebb. A dolgozat célja, hogy a variancián túl az alternatív kockázati mértékek esetén tulajdonságait feltárjuk. Tanulmányozzuk és összehasonlítjuk a következő kockázati mértékeket: szórás, abszolút

eltérés, Expected Shortfall és annak extrém verziója a maximális veszteség. Bár helyenként utalunk a Value-at-Riskre, és néhány más kockázati mértékre, melyeket a jelenlegi nemzetközi tőkeszabályozások alkalmaznak, nem fogjuk őket szisztematikusan összehasonlítani a többi mértékkel, mivel a VaR és néhány más szabályozó mérték nem felel meg a konvexitás feltételének (Artzner et. al. [1997]; Kondor et. al. [2004]). Továbbá, a szabályozó mértékek nem interpretálhatók az alapul szolgáló eloszlás funkcionáljaként sem. Végül egy megvalósíthatósági problémát tárgyalunk, miszerint a szakirodalomban erősen javasolt kockázati mértékek (pl. Expected Shortfall) fölötti optimum egy valószínűségi kérdés, amely a mintától függ. Ennek a jelenségnek hasonló megvalósíthatósági problémákhoz van köze a lineáris programozás és a véletlen geometria területén (Todd [1991]; Schmidt et. al. [1977]).

A dolgozat további részének felépítése a következő. Előbb a klasszikus kockázati mérték, a variancia zajérzékenységet vizsgáljuk. A variancia kockázati mértékként való használata feltételezi, hogy a portfólióban lévő eszközök normál valószínűségi eloszlást követnek. Amint azt már említettük, a zaj hatása és szűrési összefüggése a Gaussi portfóliókkal már számos tanulmány témája volt. Ezt szimulációs megközelítés alapján tesszük, melynél képesek vagyunk a variancia zajérzékenységének tesztelésére (Pafka et. al. [2002]). Pafka et. al. [2002] szimulációs környezetben bevezette a becsült kovariancia mátrixú és a becslési hibától mentes portfóliók szórásának hányadosát, és használta a zaj mértékének megállapítására. Ezen felül megmutatta, hogy a zaj  $N/T$  (portfólióelemszám és a mintahossz hányadosának) függvénye. Felépített egy variancia-kovariancia mátrix modellt, ami összefogja a valódi piacon megfigyelt empirikus kovariancia mátrixok lényeges tulajdonságait. Ez a modell a szakirodalomban jelenleg használt faktor modellekkel is kapcsolatba hozható. Ehhez kapcsolódóan vizsgálta a kovariancia mátrix spektrumát, és a becslési hiba arra gyakorolt hatását.

Ezt követően térünk ki az alternatív kockázati mértékekre. Az abszolút eltérést, Expected Shortfallt és ennek pesszimista verzióját, a Maximális Veszteséget (Maximal Loss, ML) vesszük itt számba. Ezen kockázati mértékeknek az elterjedt szabályozói mértékekkel (itt nem a VaR-ra, hanem a mesterségesen szabályozói súlyokon alapuló mértékekre gondolunk) közös az a tulajdonsága, hogy részben lineárisak, vagy, máshogy megfogalmazva iso-kockázati felszínük poliéder a portfóliósúlyok függvényében. Azt állítjuk, hogy ez a tulajdonság a magyarázata annak, hogy ezen kockázati mértékek nagyobb mértékben érzékenyek a zajra a

szóráshoz képest. Egy másik döntő faktor, hogy a VaR, ES és ML egy adott küszöb alatt az összes adatot elveti, azaz csak az adatok egy részének jut szerep a becslés során.

Végül áttérünk arra, hogy Expected Shortfall fölötti optimum egy valószínűségi kérdés, amely a mintától függ.

Az eredményeket továbbra is illusztrált formában, szimulációkkal alátámasztva prezentáljuk, és gyakran utalunk geometriai témákra, melyek a szakirodalomban ilyen kontextusban ritkán fedezhetük fel, ugyanakkor a megértést nagymértékben segítik.

## 11. A variancia kockázatérzékenysége

### 11.1. A Markowitz probléma

Az árak és hozamok fluktuációi egy többváltozós sztochasztikus folyamatot írnak le. Ez utóbbi együttes eloszlása legegyszerűbben többváltozós normális eloszlással modellezhető. Ez a megfontolás igen régre vezet vissza (Bachelier [1900]), ugyanakkor napjaink tipikus tankönyvi modelljei is ezen alapulnak. A piaci tapasztalatok különböző mértékben felelnek meg ennek a feltételezésnek eszköztől, likviditástól, időtávtól stb. függően. A dolgozat ezen részében mi is azt feltételezzük, hogy a hozamok együttes eloszlása többváltozós Gauss.

A racionális portfólió választás problémáját Markowitz [1952] fogalmazta meg, kapcsolatot teremtve a hozam és a kockázat között. Gaussi portfólió esetén a természetesen adódó kockázati mérték a szórás, vagy variancia, minden más ésszerű kockázati mérték szükségszerűen arányos vele. Racionális befektetők adott szintű hozam elvárása mellett, minimalizálni akarják a kockázatukat.

Matematikai kifejezésekkel élve a feladat: minimalizálni kell az alábbi kvadratikus formulát a súlyok függvényében:

$$\sigma_p^2 = \sum_{i,j=1}^N w_i \sigma_{ij} w_j \quad (28)$$

az alábbi feltételek mellett

$$\sum_{i=1}^N w_i = 1 \quad (29)$$

és

$$\sum_{i=1}^N w_i \mu_i = \mu, \quad (30)$$

ahol  $\sigma_p$  a portfólió szórása,  $\sigma_{ij}$  a kovariancia mátrix,  $w_i$  az  $i$ . eszköz súlya a portfólióban,  $\mu$  a portfólió várható hozama, míg  $\mu_i$  az  $i$ . eszköz hozamának a várható értéke. A short selling megengedett, így  $w_i$  tetszőleges előjelű lehet.

Ez a klasszikus probléma analitikusan is megoldható, lásd (Merton [1972]). (A short selling-et egy újabb lineáris feltétel megadásával zárhatjuk ki, és a feladat kvadratikus programozási feladattá válik.)

A várható hozamok megbízható becslése közismerten nehéz feladat. Mivel célunk a kockázati mértékek zajérzékenységének vizsgálata, ezért a várható hozamok tekintetében nem élünk korlátozó feltevessel. Ez azt jelenti számunkra, hogy a minimális kockázatu portfóliót kell meghatároznunk.

Első pillantásra a hozamfeltételek elhagyása értelmetlenné teszi a feladatot, de speciális esetekben (benchmarking, index tracking) éppen erre van szükség. Az analitikus megoldás tehát:

$$\omega_i^* = \frac{\sum_{j=1}^N \sigma_{ij}^{-1}}{\sum_{j,k=1}^N \sigma_{jk}^{-1}}. \quad (31)$$

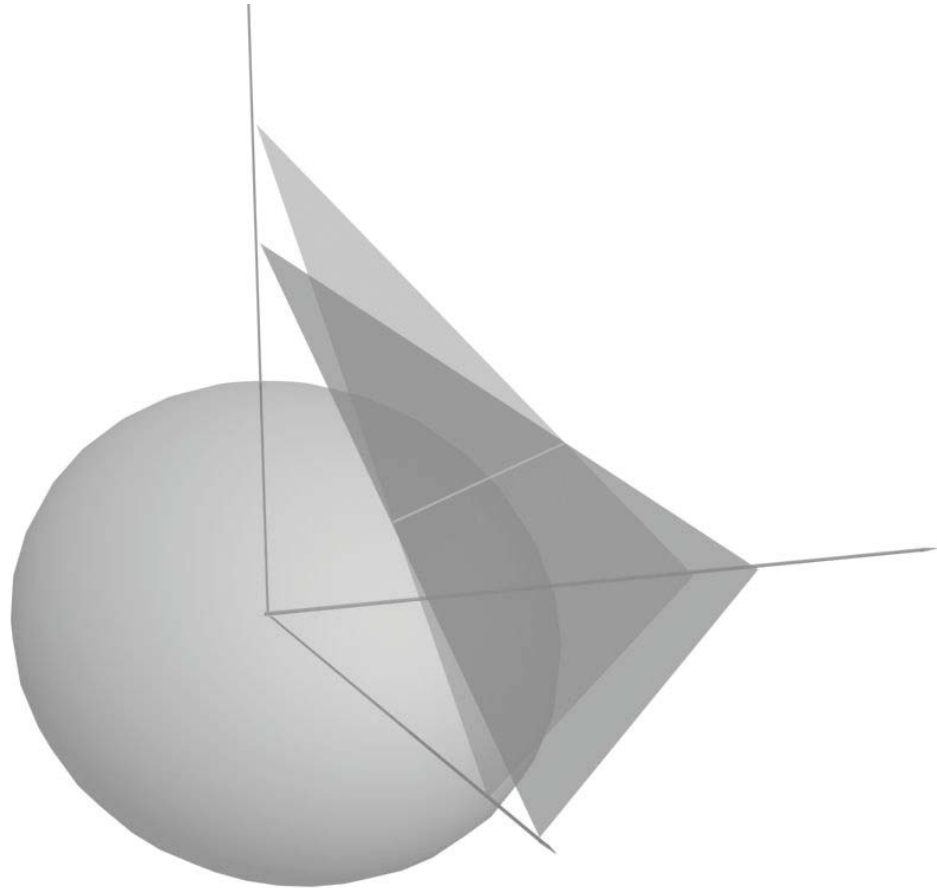
Fontos megjegyeznünk, hogy az optimális súlyok a kovariancia mátrix inverzéből adódnak. Mivel a kovariancia mátrixnak szinte szükségszerűen vannak kis sajátértékei, a mérési hibák felerősödve jelentkeznek, és a kapott portfólió érzékenyvé válik a zajra. Ez az alapvető oka annak, hogy meg kell különböztetnünk a mérési és optimalizálási problémát egymástól, előbbi esetén ugyanis nincs szükség a kovariancia mátrix inverzére.

### 11.2. A Markowitz probléma geometriai interpretációja

A Markowitz feladatnak egyértelmű geometria interpretációja van, ami segít a probléma intuitív megértésében. A kovariancia mátrix pozitív definit. Ekkor  $\sigma_p$  egy adott értékére az (28)-as egyenlet egy (hiper-) ellipszoid egyenlete. Más szavakkal  $\sigma_p$  iso-risk felületei ellipszoidok. Ezen iso-risk felület konvex volta azt a tényt tükrözi, hogy a variancia egy konvex kockázati mérték.

Az ellipszoidok főtengelei a kovariancia mátrixhoz tartozó sajátértékek gyökeinek a reciprokával arányosak. A hosszú főtengelek tehát a kis sajátértékekhez kötődnek. Ha a kockázatmentes eszközt is bevennénk portfólióba, akkor a megfelelő főtengeley végtelen hosszú lenne, és az ellipszoid hengerré alakulna. A dolgozat során nem foglalkozunk olyan portfóliókkal, melyek kockázatmentes eszközt is tartalmaznak.

A (29-30) egyenletek 1-1 hipersíkot írnak le. Geometriailag a probléma megoldásának rajta kell lennie azon a hipersíkon, amely a fenn említett hipersíkok metszetéből áll elő. Az ellipszoidot addig kell „felfűjni”, amíg az pontosan érinti a hipersíkok metszetéből előálló hipersíkot. Mivel a megoldást egy ellipszoid (konvex felület) és egy hipersík érintkezése adja, így az egyértelmű. A megoldást stabilnak nevezhetjük abból a szempontból, hogy amennyiben valamilyen  $\varepsilon$  hiba következtében az ellipszoid alakja megváltozik, akkor a megoldás helyzete folytonosan változik. A megoldásban bekövetkezett változás természetesen függ az ellipszoid görbületétől az érintési pont környezetében, ez lehet kisebb vagy nagyobb, de folytonos függvénye lesz  $\varepsilon$ -nak (arányos annak négyzetgyökével). A geometriai interpretáció a 45. ábraán látható.



45. ábra: A Markowitz feladat geometriai interpretációja  $N = 3$  esetén. A megoldást az a pont jelenti, ahol a felfűjt ellipszoid felülete érinti a korlátozó feltételek metszetéből kialakult hipersíkot.

### 11.3. Empirikus kovariancia mátrixok

A kovariancia mátrixot a  $t$  időpontban megfigyelt  $x_{it}$  hozamokból kell megbecsülnünk a  $T$  időtáv felett. Ennek a maximum likelihood becslése:

$$\sigma_{ij} = \frac{1}{T} \sum_{t=1}^T x_{it} x_{jt} \quad (32)$$

(feltéve, hogy a hozamok várható értéke 0).



Egy  $N$  elemű portfólió kovariancia mátrixa  $O(N^2)$  elemet tartalmaz. Az idősor hossza  $T$ , az adatmátrix  $N \times T$  elemet tartalmaz. Ahhoz, hogy becslésünk kellően pontos lehessen az  $N \ll T$  feltételnek kell teljesülnie, de a bankok portfóliója több száz, vagy ezer pénzügyi eszközt is tartalmazhat, míg a rendelkezésre álló (stacionárius) idősor hossza erősen függ a gazdasági környezettől.

Részvény portfóliók esetében aligha lehet ésszerű 4-5 évnél hosszabb idősort választani a gazdasági és kormányzati változások miatt, ami  $T \sim 1000$ -et jelent. Emiatt  $N/T \ll 1$  ritkán teljesül a gyakorlatban (Pafka et. al. [2002]). Ennek pedig az a következménye, hogy becslésünk hibája nagy lesz. Kellően nagy  $N$ -re és  $T$ -re azt várjuk, hogy a hiba nagysága mindössze az  $N/T$  hányadostól függjön.

A fent vázolt probléma csupán egyfajta megtestesülése a "dimenziók átkának". A közgazdászok már évek óta küzdenek ezzel a problémával (lásd pl. Gruber et. al. [1995]). Mivel a probléma gyökere abban rejlik, hogy nincs elegendő információnk, a reményt egyfelől az jelentheti, hogy kívülről viszünk információt a becslésbe. Ez számunkra annyit jelent, hogy hipotézisekkel élünk a piac struktúráját illetően, ami ugyan torzítja a becslést, de ugyanakkor bízunk benne, hogy a cserébe kapott zajcsökkenés kompenzálja ezt. Az ilyen jellegű szűrési módszereknek, bayesi becsléseknek kiterjedt irodalma van a témában.

Mivel célunk a kockázati mértékek zajérzékenységének a mérése, ezért a jelen bekezdésben említettekhez kapcsolódóan nem mutatunk be módszereket.

#### *11.4. Egy érdekes megfigyelés*

1999-ben két csoport egymástól függetlenül ugyanazt az érdekes felfedezést tette (Laloux et. al. [1999]; Plerou et. al. [1999]). Arra lettek figyelmesek, hogy a zaj hatása annyira erős a kovariancia mátrixokra, hogy a mátrix sajátértékeinek a nagy többsége (a publikált példában azok 94 százaléka) a véletlen mátrixokra jellemző spektrum tartományába esik. A véletlen mátrixok nagy érdeklődést váltottak ki elsősorban a fizikusok körében, ezért Laloux et. al. [2000] egy szűrési eljárást javasolt a véletlen mátrixok elméletére (RMT) alapozva. Ezt többen továbbfejlesztették, és finomították abban bízva, hogy a véletlen spektrumra jellemző tartományból további információkat lehet kiszűrni (Plerou et. al. [2002]; Jurkiewicz [2004]; Papp et. al. [2005]).

### 11.5. Wishart mátrixok sajátérték eloszlása

Hogy megérthessük a javasolt szűrési eljárás lényegét, felvázolunk néhány egyszerű tényt a véletlen mátrixokkal kapcsolatban.

Legyen  $X$  egy  $N \times T$  adatmátrix, melynek elemei  $x_{it}$  független azonos eloszlású (i.i.d.) véletlen változók nulla várható értékkel és véges második momentummal. Ekkor, ha  $N, T \rightarrow \infty$ ,  $N/T < 1$  rögzített értéke mellett, a  $\sigma = \frac{1}{T} X X'$  mátrix (ahol  $X'$  a transzponáltat jelöli) sajátértékeinek az eloszlása ( $\rho(\lambda)$ ) a Wishart, vagy más néven a Marchenko-Pastur spektrumhoz konvergál (Marchenko et. al. [1967]):

$$\rho(\lambda) = \frac{1}{N} \frac{dn(\lambda)}{d\lambda} = \frac{T}{2\pi N} \frac{\sqrt{(\lambda - \lambda_{\min})(\lambda_{\max} - \lambda)}}{\lambda}, \quad (33)$$

ahol

$$\lambda_{\max, \min} = 1 \pm \sqrt{\frac{N}{T}} \quad (34)$$

Néhány megjegyzés a fentiekhez kapcsolódóan:

- Ha  $X$   $N \times T$ -s adatmátrix, akkor  $\sigma = \frac{1}{T} X X'$  pontosan a minta kovariancia mátrixa.
- Ha  $T < N$ , akkor a  $1 - T/N$  súllyal nulla sajátértékek jelennek meg.
- Ha  $T = N$ , akkor Marchenko-Pasteur tétele összefüggésben van a Wigner féle félkör tétellel (Wigner [1958]), ami az első határeloszlás tétel véletlen (négyzetes) mátrixokra.
- A konvergencia gyors, vélhetően  $\sim 1/N$  (Bai [1993]).

Ez utóbbi megjegyzés jelentősége, hogy a véges Wishart mátrixok sajátértékei is a Marchenko-Pasteur tétel szerint viselkednek, és azok többsége már  $N$  igen alacsony

### *11.6. Egy szűrési eljárás a véletlen mátrixok elméletére alapozva*

Most pedig visszatérünk arra a szűrési eljárásra, melyet Laloux et. al. [1999] javasolt. Észrevették, hogy a nagy empirikus kovariancia mátrixok spektruma a kis sajátértékeknek egy kvázi folytonos sávját tartalmazza, plusz néhány (~ 20) diszkrét közepes nagyságú sajátértéket, valamint egy darab nagy sajátértéket távol a többitől. Ezzel a nagy sajátértékekkel jellemezték a piacot, a közepes sajátértékeket a főbb ipari szektorokkal hozták kapcsolatba, míg a kvázi folytonos sáv struktúrája megkülönböztethetetlen volt a véletlen mátrixok kovariancia mátrixára jellemző spektrumtól. Ebből azt a következtetést vonták le, hogy ez a sáv nem tartalmaz információt a piac struktúrájára vonatkozóan.

A javasolt szűrési eljárás azon az ötleten alapszik, hogy ezt a sávot vegyük ki a spektrumból, és csupán a sávon túli sajátértékeket vegyük figyelembe. Információt ugyanis csak ezek a sajátértékek és a hozzájuk tartozó sajátvektorok adnak. Gyakorlatban ez azt jelenti, hogy az optimalizációt a nagy sajátértékekhez tartozó sajátvektorok altéren kell végrehajtani, a kis sajátértékeket konstanssal kell helyettesíteni, úgy hogy a mátrix nyomát (trace) megőrizzük. Ez az eljárás nagy mértékben csökkenti a probléma dimenzionalitását, és az optimalizálás jobb portfóliókhöz vezet (Laloux et. al. [2000], Plerou et. al. [2002]).

A módszer a főkomponens analízis egy fajtájának tekinthető, ahol a véletlen sáv felső értéke határozza meg a figyelembe veendő sajátértékeket és sajátvektorokat, ezáltal objektív kritériumot szolgáltatva a figyelembe veendő főkomponensek számára.

Pafka et. al. [2004] elvégzett egy összehasonlítási tesztet, s azt találta, hogy ez a szűrési eljárás valóban jól működik. Arra, hogy a véletlen sávból további információkat tudjunk kiszűrni (Jurkiewicz [2004] és Papp et. al. [2005]) javasolt eljárásokat.

Habár a véletlen mátrixok elméletén alapuló szűrési eljárás versenyképesnek bizonyul, és gyakran jobb portfólióhoz vezet, mint más szűrési eljárások, néhány kérdés így is felmerül.

- Ha az empirikus kovariancia mátrix leginkább csak zajból áll, hogy lehet az, hogy a pénzügyi intézmények mégis széleskörűen használják?
- A nagy számú, véletlen sávba eső sajátértékek milyen mértékig jellemzik jól a zaj hatását a portfólió minőségére? Tegyük föl, hogy az optimalizálandó feladat megoldása a legnagyobb sajátértékekhez tartozó sajátvektorok alterében fekszik, és ezek mintáról mintára viszonylagos stabilitást mutatnak. Ez esetben sem képzelhető el, hogy a véletlen sajátértékek alterében még a lényeges fluktuációk is csak csekély hatást gyakorolnak az optimális portfólióra?
- Ráadásul Laloux et. al. [2000] a többiekhez hasonlóan, akik a szűrési problémával foglalkoztak, valós empirikus adatokat használtak, ahol további nem kívánt hatások (mint például az idősor nem stacionárius jellege) lehettek jelen, és tovább növelhették a zaj forrását, együttesen azzal az alapvető nehézséggel, amelyet az idősor véges hossza okozott.

Ezen kérdésektől indítva úgy döntöttünk, hogy laboratóriumi körülmények között vizsgáljuk meg a problémát. Ahelyett, hogy valós adatokat használnánk, mesterséges idősorokat generáltunk, melyben teljes kontrollal rendelkezünk a mögöttes sztochasztikus folyamatról. Ezt az motiválta, hogy képesek legyünk összehasonlítani a különböző kockázati mértékek zajérzékenységét, és ehhez jobb, ha megszabadulunk a bizonytalanság más forrásaitól.

Stratégiánk az volt, hogy különféle kovariancia mátrix modelleket használtunk, és ezekből idősorokat generáltunk. Ezt követően egy  $T$  hosszúságú szakaszt kivettünk az idősorból, mintha azt a piacon figyeltük volna meg, és megpróbáltuk belőle visszaállítani a kovariancia mátrixot. Ezután optimalizáltuk a portfóliót mind az igazi, mind a megfigyelt kovariancia mátrixból adott kockázati mérték mellett, és meghatároztuk az idősor véges voltából adódó hiba nagyságát. Ezeket a modelleket úgy választottuk meg, hogy valamilyen mértékben jellemezzék a valós piacot. Négy ilyen kovariancia mátrix struktúrát mutatunk be, komplexitásuknak megfelelő sorrendben.

### 11.7. Az optimális portfólióra gyakorolt zaj hatásának mérése

A zaj hatásának a mérésére különféle metrikákat használhatnánk, a kovariancia mátrixok, vagy akár ezen mátrixok sajátértékei között különféle távolságokat vehetnénk. Ezzel szemben azt gondoljuk, hogy egy portfólió szuboptimalitását azzal jellemezhetjük leginkább, hogy a zaj hatására mennyivel növekedett meg  $(q_0 - 1)$  a portfólió kockázata, ahol

$$q_0^2 = \frac{\sum_{ij} w_i^* \sigma_{ij}^{(0)} w_j^*}{\sum_{ij} w_i^{(0)*} \sigma_{ij}^{(0)} w_j^{(0)*}}. \quad (35)$$

Itt  $\sigma^{(0)}$  az igazi kovariancia mátrix, a hozzá tartozó portfóliósúlyokat  $\omega^{(0)*}$  jelöli.  $\omega^*$  jelöli az empirikus mintából adódó kovariancia mátrix fölött végzett optimalizálás eredményét, míg a hozzá tartozó portfólió kockázatát a (35)-ös képlet számlálója tartalmazza. Ezzel szemben az igaz portfólió kockázata a (35)-ös képlet nevezőjében található.

Ezt a mértéket Pafka et. al. [2002] vezette be. Ez természetesen azt feltételezi, hogy ismerjük a valódi generáló folyamatot.

A (8)-as egyenlet természetesen egy adott minta fölött érvényes, azaz más  $T$  időintervallumot véve annak értékére is mást kapunk. Más szavakkal  $q_0$  maga is egy valószínűségi változó.

### 11.8. Modell építés és eredmények

#### 11.8.1. 1. modell: A legegyszerűbb kovariancia mátrix

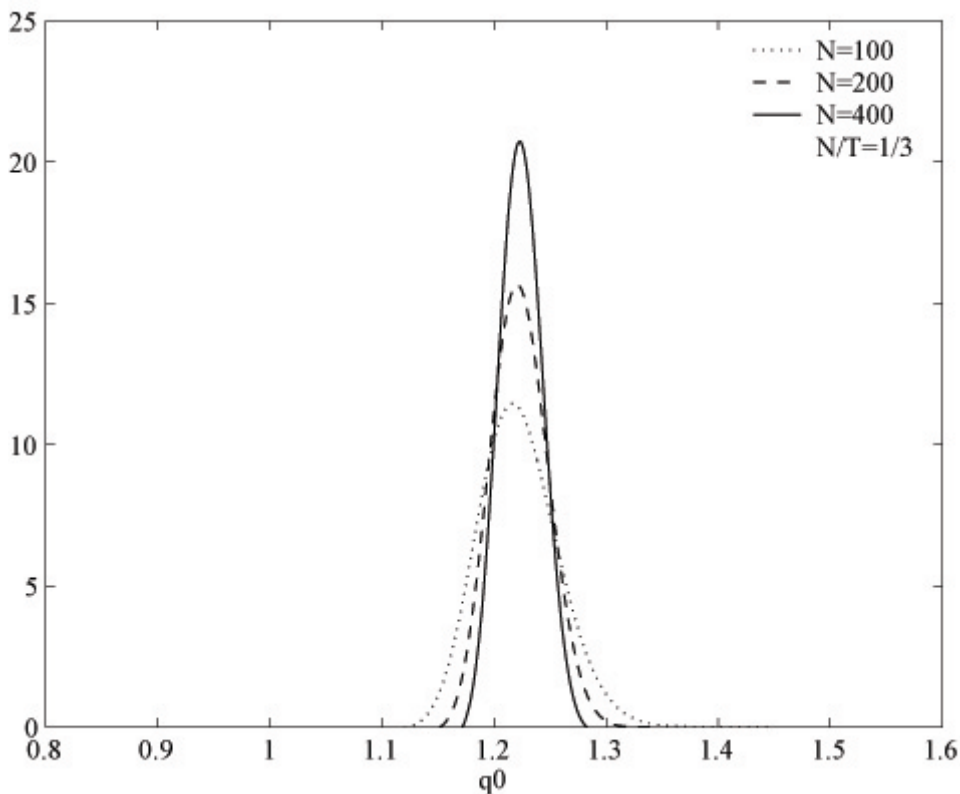
Tegyük fel, hogy a hozamok együttes eloszlása többváltozós normális, ennek is a legegyszerűbb formája, amikor az egység mátrix a kovariancia mátrix. Ez azt jelenti, hogy a kovariancia mátrixnak egy db  $N$ -szeresen degenerált sajátértéke van,  $\lambda = 1$ . Ha most generálunk egy  $N \times T$ -s adatmátrixot, és megpróbáljuk rekonstruálni a kovariancia mátrixot az (32)-es egyenletnek megfelelően, akkor egy sokkal bonyolultabb struktúrát kapunk, mint az egység mátrix. Ez a struktúra mintáról mintára változik, és konvergál az egység mátrixhoz, ha  $N$  rögzített értékére  $T \rightarrow \infty$ .  $N$  és  $T$  nagy értékeire  $N/T$  fixen tartása mellett azt tapasztaljuk, hogy a zaj hatására a véges  $T$  következtében a  $\lambda = 1$  sajátértékek felhasadnak, és a keletkezett spektrum a Marchenko Pastur spektrumhoz konvergál (lásd (33)-as egyenlet).

Pafka et. al. [2003], Nowak et. al. [2003] és Papp et. al. [2005] megmutatták, hogy  $N, T \rightarrow \infty$  esetén, és rögzített  $N/T < 1$  mellett fennáll az alábbi összefüggés:

$$q_0 = \frac{1}{\sqrt{1 - \frac{N}{T}}}. \quad (36)$$

Jól ismert, hogy az empirikus kovariancia mátrix maximális rangja  $N$  és  $T$  közül a kisebbik lesz.  $T < N$  esetén a kovariancia mátrixnak nulla sajátértékei lesznek, és a portfólió optimalizálási feladat értelmetlenné válik. A (36)-os egyenlet jól mutatja, hogy  $N/T \rightarrow 1$  esetén a relatív kockázat divergál.

(36) csak  $N; T \rightarrow \infty$  esetén érvényes. Ha  $q_0$  eloszlását szeretnénk tanulmányozni,

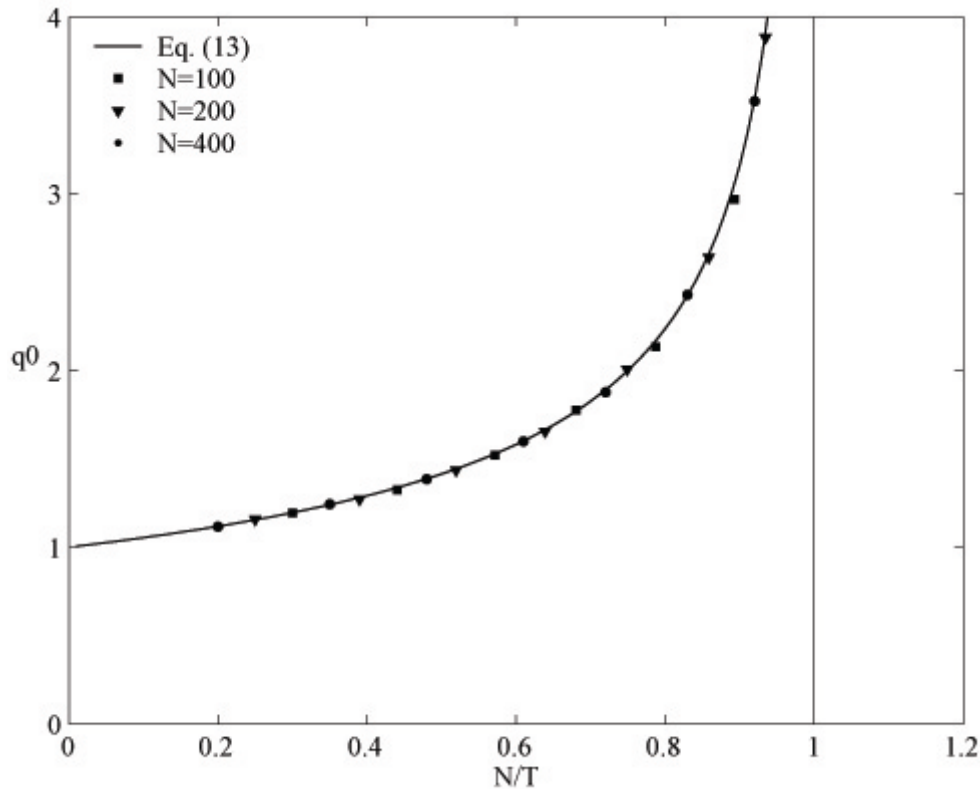


46. ábra.  $q_0$  eloszlása növekvő  $N$ -re,  $N/T$  rögzített értéke mellett (szimulációs eredmények).

akkor szimulációs eljárásokhoz kell folyamodnunk. A 46. ábra  $q_0$  hisztogramját mutatja különböző  $N$  értékekre,  $N/T$  rögzített értéke mellett. Jól látható, hogy  $N$  növekedésével az eloszlás egyre élesebbé válik. Másrészt a 47. ábra mutatja  $q_0$  átlagos értékeit különböző  $N$  és  $T$  értékekre. A szimulációból nyert adatpontok szépen illeszkednek az analitikus függvényre (ahol  $N, T \rightarrow \infty$  volt érvényben) már mérsékelt  $N$  és  $T$  értékek esetén is, illetve jól látható az  $N/T$ -től való függés is.

### 11.8.2. 2. modell: Az egyfaktoros (piaci) modell

A 2. modell egy elmozdulás a valóság felé: azt feltételezzük, hogy a véletlen változóink nem függetlenek, hanem egyetlen faktoron keresztül korrelálnak, ami a piac egészét jelenti. A hozzá tartozó kovariancia mátrixot a 48. ábra ábra tartalmazza. (Az egyszerűség kedvéért a szimmetria magas fokát feltételezzük, azaz minden diagonális elem értéke 1, és minden diagonálison kívüli elem ( $\rho_0$ )).



47. ábra:  $q_0$  átlagos értéke  $N/T$  függvényében,  $N$  különböző értékei mellett (szimulációs eredmények). Az adatok szépen illeszkednek a Eq. (13)-as egyenlettel leírt függvényre.

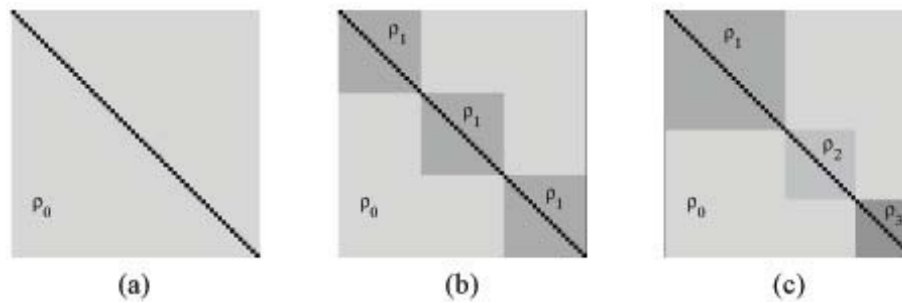
Ebben az esetben a különbség a korrelációs és a kovariancia mátrixok között megszűnik.)

Egy ilyen mátrixnak egy darab  $(N-1)$ -szeres degenerált kis sajátértéke van, melynek értéke  $1-\rho_0$ , és 1 darab nagy (Frobenius-Perron) sajátértéke, melynek értéke  $1 + (N-1)\rho_0$ . Az empirikus kovariancia mátrixban a zaj nagysága dönti el a kis sajátértékek degeneráltságát, és bepaszírozza őket a Marchenko-Pastur spektrumba, míg a nagy sajátértékre a hatása mindössze  $O(1/N)$ .

Mivel a legnagyobb sajátérték súlya a spektrál eloszlásban 1, míg a kis sajátértékek súlya  $N-1$ , ezért  $q_0$   $O(1/N)$  hibával határozható meg, így ugyanazt kapjuk, mint az 1-es modellben, lásd (36)-os egyenlet.

Szimulációk erre, és az ezt követő modellekre a következő standard eljárással készíthetők. Ha adott egy  $\sigma^{(0)}$  kovariancia mátrix, akkor az ehhez tartozó





48. ábra: A 2-es, 3-as és 4-es modell korrelációs struktúrái.

empirikus korrelációs mátrix a következőképpen állítható elő: generáljunk az igazi  $\sigma^{(0)}$  korrelációs mátrixból véges hosszú idősort,  $x_{it} = \sum_j L_{ij} z_{jt}$ , ahol  $L$  az igazi korrelációs mátrix Cholesky-felbontása ( $(\sigma^{(0)} = LL')$ ), és  $z_{jt}$  független eloszlású standard normális változó. Ekkor az empirikus korrelációs mátrix ((32)-es egyenlet) szokásos becslése a következő:  $\sigma = \frac{1}{T} \sum_{t=1}^T x_{it} x_{jt}$ .

Az 1-es táblázat a 2-es modell néhány szimulációját mutatja. Az is jól látható, hogy nagy  $N$ -re és  $T$ -re a (36)-os egyenlet egy jó közelítése  $q_0$ -nak a 3-as modellben is.

### 11.8.3. 3. Modell: Szimmetrikus piac-plusz-szektorok modell

Itt azt feltételezzük, hogy az értékpapírok szektorokba csoportosíthatók, melyek mérete  $N_1$ , és a korreláció a szektoron belül ( $\rho_1$ ), ami nagyobb, mint a szektorok közötti korreláció (48. ábra). A spektrum három különböző sajátértéket tartalmaz, egy  $(N - \frac{N}{N_1})$ -

szeres degenerált kis sajátértéket, melynek értéke  $1 - \rho_1$ , egy darab  $(\frac{N}{N_1} - 1)$ -szeres

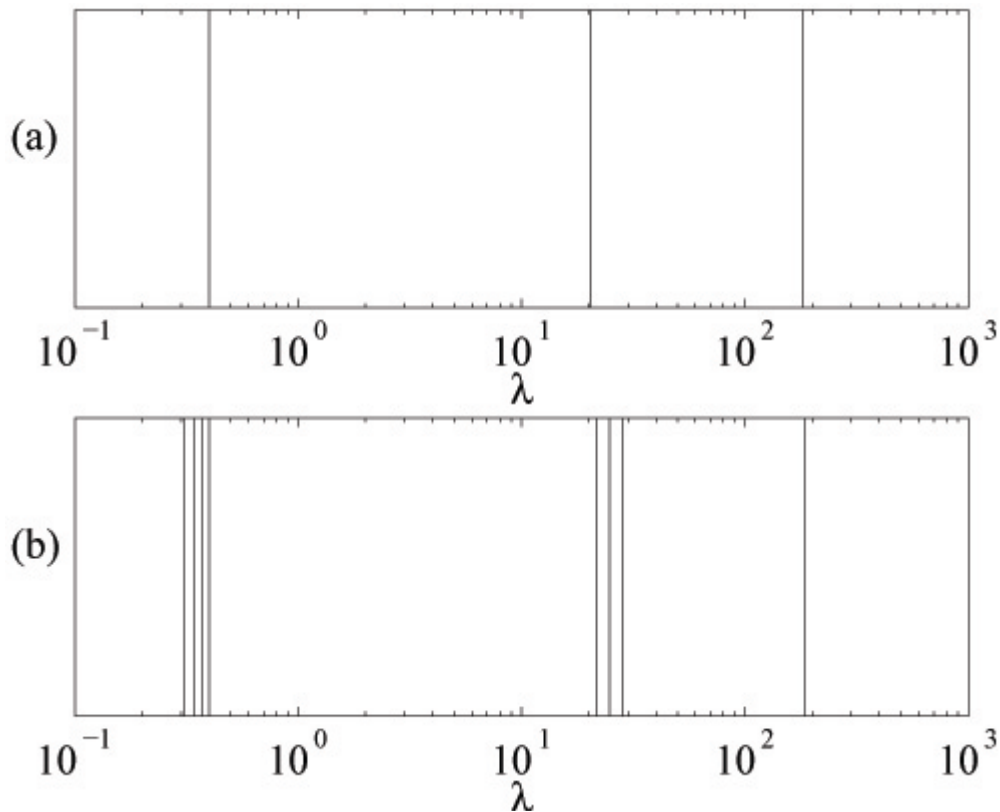
közepes sajátértéket, melyek értéke  $1 + (N_1 - 1)\rho_1 - N_1\rho_0$  és egyetlen nagy sajátértéket  $1 + (N_1 - 1)\rho_1 + (N - N_1)\rho_0$  értékkel. Ha  $N \gg N_1 \gg 1$ , akkor ezek tisztán elválnak, lásd 49. ábra, míg a zaj hatására ezek két elkülönült blokkban jelennek meg, és ezeken túl található a legnagyobb sajátérték.

Mod.	$N$	$T$	$\rho_0$	$N_k$	$\rho_k$	$q_0$ átlag	$q_0$ szórás	$q_0$ diff.	$q_0$ hiba
1	100	200				1,4103	0,0690	-0,0039	0,0022
2	100	200	0,2			1,4114	0,0700	-0,0028	0,0022
2	100	200	0,5			1,4101	0,0701	-0,0041	0,0022
3	100	200	0,2	20	0,4	1,4084	0,0719	-0,0059	0,0023
4	100	200	0,3	25	0,4-0,8	1,4117	0,0724	-0,0025	0,0023
4	100	200	0,2	5-50	0,4-0,9	1,4092	0,0703	-0,0050	0,0022
2	20	40	0,2			1,3894	0,1513	-0,0247	0,0047
3	20	40	0,1	5	0,3	1,3844	0,1511	-0,0298	0,0047
2	400	800	0,2			1,4125	0,0352	-0,0016	0,0011
4	400	800	0,3	100	0,4-0,8,	1,4126	0,0360	-0,0016	0,0011
4	400	800	0,2	50-200	0,3-0,9	1,4141	0,0354	-0,0001	0,0011

32. táblázat: Szimulációkból származó eredmények, adott paraméterek mellett. „ $q_0$  átlag” és „ $q_0$  Szórás” jelentik  $q_0$  1000 Monte Carlo szimulációból számított átlagát és szórását. „ $q_0$  diff” jelenti a különbséget „ $q_0$  átlag” és (13) -mas egyenlet között, míg „ $q_0$  hiba” a „ $q_0$  átlag” becslési hibáját mutatja.

$N_1$ ,  $\rho_0$ ,  $\rho_1$  paraméterektől függően a modell viselkedése bonyolulttá válhat, de nagy  $N$  és  $N_1$  értékekre  $q_0$  ismételten belesimul a (13)-mas egyenletbe (lásd az 32. táblázat szimulációs eredményeit).

A véletlen mátrix elméleten alapuló szűrési eljárás erős voltát Pafka et. al. [2004]



49. ábra. A sajátértékek szétválása a 3-as és a 4-es modellben.

vizsgálta, és megmutatta, hogy a módszer igen hatékonyan működik. Nem csupán a hibát csökkenti, ahogy  $N$  tart  $T$ -hez, hanem lehetővé teszi, hogy  $T$ -nél nagyobb elemszámú portfóliókat optimalizáljunk, lényegében ugyanakkora hiba mellett.

#### 11.8.4. 4. modell: Antiszimmetrikus piac-plusz-szektorok modell

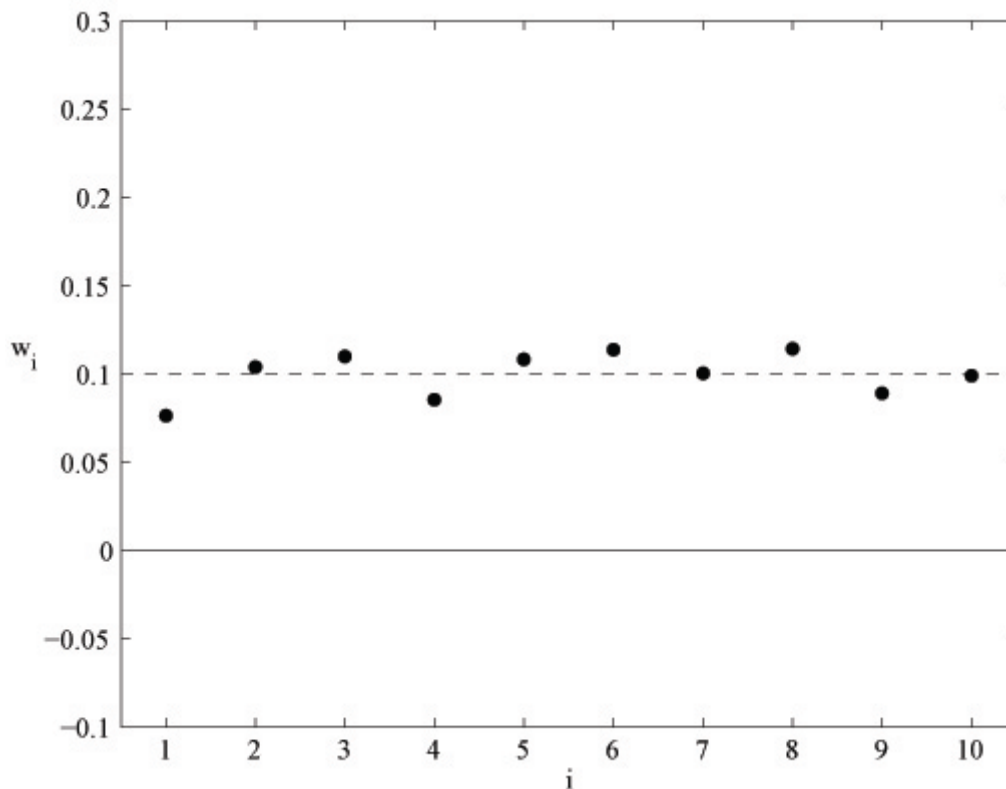
Újabb lépésként a valós piaci struktúra felé feloldjuk a szektorok közötti szimetriára tett feltevésünket. A szektoron kívüli korreláció továbbra is  $\rho_0 > 0$ , míg a szektoron belüli korreláció  $\rho_k$ , ( $k = 1 \dots K$ ). A szektorokon belüli korrelációk a diagonális blokkokban különbözőek, de blokkonként ugyanazok (és nagyobbak, mint a blokkon kívüli korreláció,  $\rho_k > \rho_0$ ). A blokk méretére ( $N_k, k = 1 \dots K$ ), vonatkozóan is megengedjük, hogy azok különböző nagyságúak legyenek. A modell, amit most leírunk, megegyezik azzal, amit Noh [2000] bevezetett. Az egyszerűség kedvéért továbbra is fenntartjuk azt a feltevésünket, hogy a korrelációs és a kovariancia mátrix megegyezik, azaz az egyes instrumentumok varianciáját 1-re állítottuk be. A korrelációs mátrix struktúráját a 4 (c) ábra mutatja.

Egy ilyen mátrix, mely  $K$  szektort tartalmaz,  $K$  db kis sajátértékkel rendelkezik, melyek értéke:  $1 - \rho_k < 1$ ,  $k = 1 \dots K$ . A megfelelő sajátvektoroknak mindössze 2 nem nulla eleme van, melyek abszolút értékben megegyeznek. Multiplicitásuk  $N_k - 1$ ,  $k = 1 \dots K$ , ahol  $N_k$  a  $k$ . szektor mérete, azaz a teljes multiplicitása a kis sajátértékeknek  $N - K$ . Továbbá,  $K$  db nagyobb sajátérték létezik ( $\lambda > 1$ ), tipikusan egyedül állnak és a modell minden paraméterétől ( $\rho_0, \rho_k, N_k$ ) függenek. Ez egy  $K$  szektoros mátrix,  $2K$  db különböző sajátértékkel. A Frobenius-Perron tétel alapján a legnagyobb sajátérték szükségszerűen elkülönül, és a hozzá tartozó sajátvektor minden komponense pozitív. Ez az érték a teljes piaccal azonosítható, míg a többi  $K - 1$  sajátérték a szektorokhoz köthető, és megfelel az előző modell közepes sajátértékeinek. Ez alkalommal ezek degeneráltsága a szimmetria hiányán (blokkok különböző voltán) múlik (49. ábra. A sajátértékek szétválása a 3-as és a 4-es modellben.).

Az adott empirikus mátrix sajátérték struktúráját a korábbiakhoz hasonlónak találjuk, a kis sajátértékek felhasadnak, és a már jól ismert Marchenko-Pastur spektrum adódik. A legnagyobb sajátérték alapvetően érzéketlen a zajra, míg annak hatása a közepes sajátértékekre a paraméterek függvénye. Ha a paraméterek olyanok, hogy ezek a sajátértékek jól elkülönülnek és különbözőek, továbbá  $N/T$  kellően kicsi, akkor a közepes sajátértékek különállóak maradnak az empirikus spektrumban is.

Összességében ez az egyszerű modell reprodukálja az empirikus kovariancia mátrixok főbb jellemvonásait a spektrumban (Laloux et. al. [1999]; Plerou [2002]).

Szimulációk (lásd 32. táblázat eredményei) azt mutatják, hogy nagy  $N$ -re és  $T$ -re  $q_0$  jól közelíti a 13-mas egyenletet a 4-es modellben is.



50. ábra A becsült korrelációs mátrixból származó „optimális”  $w_i^*$  ( $i=1..N$ ) portfóliósúlyok, míg a szaggatott vonal az igazi súlyokat mutatja viszonyításképpen  $N = 10, T = 500$  esetén.

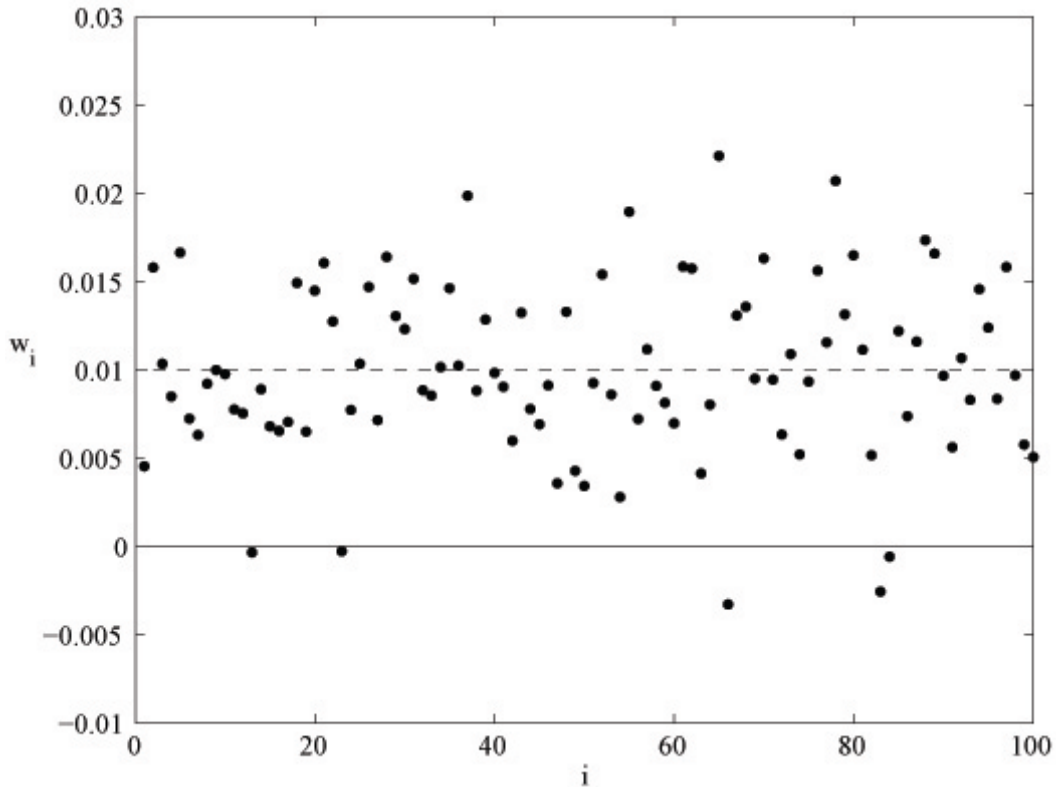
### 1.9. A súlyok ingadozása és $q_0$ eloszlása

Láthattuk, hogy a zaj hatása az átlagos relatív standard hibára esetenként igen jelentős, különösen nagy portfóliók és nem elégséges idősor esetén. Az átlagos  $q_0$  mindössze a zaj nagyságának egy globális mérőszáma, részletesebb jellemzése az empirikus kovariancia mátrix optimalizálása során adódó súlyok megfigyelésével adható meg. Ahogy a mátrix mintáról mintára változik, a súlyok is úgy változnak.

A 50. ábraán egy kis portfólió ( $N = 10$ ) optimális súlyai láthatók, míg a 51. ábraán ugyanezt egy nagyobb portfólió esetében ( $N = 100$ ) láthatjuk.

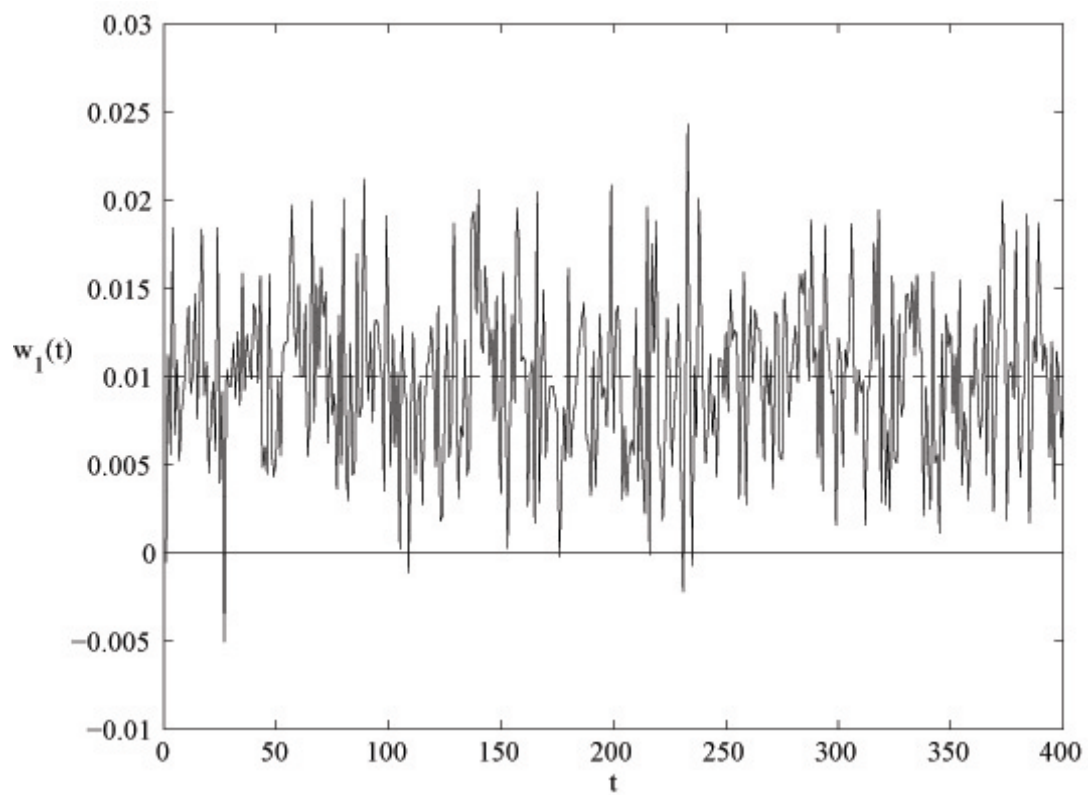
Az eszközök hozamai független standard normális eloszlásúak. A szimuláció eredményei egy  $T = 500$  hosszú időorból származnak, azaz az  $N/T$  hányados kellően messze van a kritikus  $N/T = 1$  küszöbtől. Ebben a példában az eszközöket teljesen ekvivalensnek tekintjük, így az igazi súlyok rendre  $1/N$  nagyságúak, és egyenlőek. Ezek

az ábrán is láthatók összehasonlításképpen. Az empirikus súlyoknak az igazi súlyoktól való távolsága azonban már meglepő.

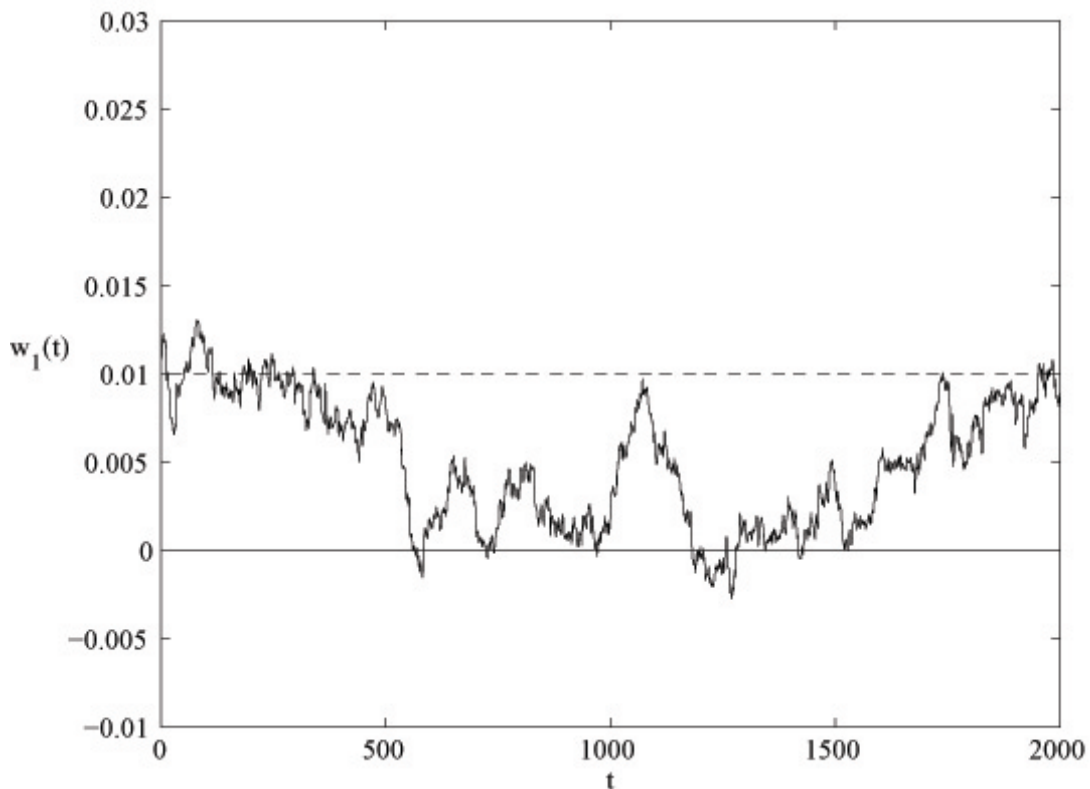


51. ábra: A becült korrelációs mátrixból származó „optimális”  $\omega_i^*$  ( $i = 1 \dots N$ ) portfóliósúlyok, míg a szaggatott vonal az igazi súlyokat mutatja viszonyításképpen  $N = 100$ ,  $T = 500$  esetén.

Az 52. ábra az empirikus súlyok ingadozását mutatja időben,  $N = 100$  és  $T = 500$  esetre. A 8-as ábrán egy adott súly alakulását mutatjuk be, melyet  $T$  hosszúságú, és egymást át nem fedő idősorokból kaptunk, míg a 9-es ábrán ugyanez látható azzal a különbséggel, hogy az idősort mindig 1 időegységgel toltuk előre. Az első esetben a súlyok vad fluktuációja jellemző, míg a másodikban az egymást követő  $T$  hosszúságú idősorok egymást átfedők, így az azokból származó súlyok erősen korreláltak lesznek, ami a súly időbeni alakulását leíró trajektóriáját simábbá teszi, de ugyanakkor távol marad az igazi értékétől. Nyilvánvalóan sem az egyik, sem a másik eset nem túl ígéretes a portfólió összeállítása szempontjából. Az első esetben a portfóliót teljesen át kell rendezni minden  $T$  periódusban, míg az utóbbi esetén a portfólió stabil marad egyik napról a másikra, de az idő nagy részében távol marad a valóban optimális összeállítástól.



52. ábra: A  $w_1^*$  súly alakulása  $T = 500$  hosszúságú egymást nem átfedő idősorokból  $N = 100$  esetén (a szaggatott vonalak az „igazi súlyokat jelzik”)



53. ábra: A  $\omega_1^*$  súly alakulása  $T = 500$  hosszúságú egymást átfedő idősorokból  $N = 100$  esetén (a szaggatott vonalak az „igazi súlyokat jelzik”)

A súlyok nagyfokú instabilitásához viszonyítva a portfólió kockázatának az ingadozása relatíve kicsi. Emlékezzünk a 46. ábraára, amely  $q_0$  eloszlását mutatja, ahogy az mintáról mintára változik. Nagy  $N$ -re és  $N/T$  kellően kicsi értékére az eloszlás éles, és bármely mintán kapott érték megfelelően reprezentálja az átlagot. Ha  $N/T$  1-hez közele, akkor az eloszlás értelmezési tartománya és a súlyok varianciája divergál.

Minden eredmény, amely szimulációs modellünk eredménye kétséget ébreszthet az olvasóban az eredmények relevanciájának a tekintetében. Ezek az eredmények ideális körülmények között keletkeztek, a generált hozamok normális eloszlásúak, és az idősor erősen stacionárius. Ezek egyike sem jellemző a valós piacokra, de pontosan emiatt azt gondoljuk, hogy a zaj hatása valós portfóliókon még erősebb lenne, mint egy ilyen idealizált környezetben.

Figyelembe véve azt a tényt, hogy  $N/T$  sosem kicsi a gyakorlatban, sőt gyakran a kritikus  $N/T = 1$  alá esik, szükségeszerű, hogy valamilyen szűrési eljárást alkalmazzunk a



zaj csökkentése érdekében. Ezen technikák nagy része megtalálható a szakirodalomban (Gruber et. al. [1995]), és széles körben használt. A szűrési módszerek ismertetése nem célunk ebben a tanulmányban. A következő fejezetben bemutatjuk néhány alternatív kockázati mértékek zajérzékenységét, melyekre már nem léteznek szűrési technikák. Hogy azok zajérzékenységét össze tudjuk hasonlítani a varianciáéval, ezért ez utóbbinak a szűretlen esetre vonatkozó eredményeit vesszük figyelembe.

## *12. Alternatív kockázati mértékek*

Mivel a normális eloszlásnak két el nem tűnő kumulánsa van, a várható érték és a variancia, így a fluktuációk tipikus nagysága egy számmal jellemezhető, a varianciával, vagy a szórással. Más eloszlások esetén a variancia nem elegendő. Vastag szélű eloszlások esetén, melyek aszimptotikusan hatványfüggvényszerűen csökkennek, a variancia egy félrevezető kockázati mérték. A valós portfóliók gyakran rendelkeznek ezzel a tulajdonsággal, ezért nem tanácsos őket variancia mellett minimalizálni. Ezt felismerve az irodalom és a gyakorlat számos alternatív kockázati mértéket kínál. Nem kívánjuk ezek mindegyikét felsorolni, csupán néhány gyakorlati és elméleti jelentőségűt vizsgálunk.

### *12.1. Value at Risk és a szabályozói mértékek*

A Value at Risk (VaR) és más szabályozói mértékek nem tárgyai dolgozatunk ezen részének. A VaR-t itt újra bemutatjuk, ugyanakkor a korábbiaktól eltérő aspektusban, mivel fontosnak tartjuk tisztázni, hogy miért nem illik ezen fejezet kereteibe.

A Value at Risk, mint a dolgozat korábbi fejezeteiben bemutattuk nem más, mint egy általában magas konfidenciaszinthez (95%, vagy 99%) tartozó kvantilis, egy küszöb. 1994-ben vezette be a Riskmetrics, és gyorsan elterjedt az egész bankszektorban, és beépült a szabályozásba is. Kétségtelenül vannak megkérdőjelezhetetlen előnyei a „mi lenne ha” típusú kockázati mértékekkel szemben (például görögök): univerzális, ami azt jelenti, hogy tetszőleges portfólióra alkalmazható, a sűrűségfüggvény funkcionálja,

valószínűségi tartalommal bír és pénzben van kifejezve. Vannak azonban hátrányai is, ezek közül a legnyilvánvalóbb, hogy nem mond semmit a veszteség átlagos, tipikus nagyságáról. Másrészt mivel kvantilisról beszélünk, nem egy konvex kockázati mérték. Ez különösen azért zavaró, mert az elmúlt években kockázatomérő eszközökből döntéshozó eszközzé nőtte ki magát, ami esetünkben portfólió optimalizálást jelent, amit különösen nehéz elvégezni egy nem konvex kockázati mérték mellett. (A VaR előnyeiről és hátrányairól többek között az alábbi szerzőknél olvashatunk: Shin et. al. [2001], Acerbi et. al. [2002], McNeil [2002], Rockafellar et. al. [2002].

A legtöbb szerző számára a konvexitás hiánya volt a fő motiváció, hogy új kockázati mértéket keressen, és ez vezetett a koherens kockázati mértékek axiómáinak a lefektetéséhez. (Artzner et. al. [1997]; Heath et. al. [1999]). Az axiómarendszer legfontosabb jellemzőit Gáll et. al. [2005] alapján foglaljuk össze. Egy kockázati mérték akkor koherens, ha az alábbi tulajdonságokat teljesíti.

- Monoton: ha egy portfólió minden esetben nagyobb hozamot ígér, mint egy másik, akkor annak ne legyen nagyobb a kockázata.
- Pozitív homogén: Megtöbbszörözve a portfóliót, miközben annak összetételét megtartjuk, a portfólió kockázata is arányosan növekedjen.
- Szubadditív: Két portfóliót összetéve ne növekedhessen a kockázat abban az értelemben, hogy az egyes portfóliók kockázatának összegét nem haladhatjuk meg. Ebben a tulajdonságban tükröződik vissza a diverzifikáció fogalma.
- Eltolás invariáns: Ha biztosan realizálunk egy adott összegű pénzáramlást, akkor a portfólió kockázata épp ennek a pénzáramlásnak a mértékével csökkenjen.

A fenti tulajdonságok tükrében, ha egy mérték pozitív homogén és szubadditív, akkor az konvex is, az állítás megfordítva ugyanakkor nem igaz. A konvexitás elvárása egy gyengébb követelmény, mint külön-külön elvárni egy mértéktől, hogy pozitív homogén, és szubadditív is legyen. Ha azonban egy kockázati mérték a szubadditivitás tulajdonságának nem tesz eleget, akkor az nem is konvex. Ez fordítva is igaz, ha mérték nem konvex, akkor nem is szubadditív. Ebből adódóan a dolgozat során kvázi szinonimaként használjuk a konvexitás és szubadditivitás meglétét, vagy hiányát, holott nem ugyanazon, de szorosan összefüggő fogalmakról van szó.

A konvexitás hiányából eredő hátrányok nagy portfóliók esetén vitathatók, jelen tanulmányunkban a VaR-t nem kívánjuk vizsgálni, a bonyolult, nem konvex viselkedése miatt, ami alacsony  $N/T$  arány mellett is jelen van, s ezzel lehetetlenné teszi a közvetlen összehasonlítást a varianciával.

A VaR-t a nemzetközi szabályozás karolta föl, azáltal hogy a Basell-es tőkeegyezmény az 1996-os kiegészítésében lehetővé tette a bankok számára, hogy kereskedési könyvük tőkekövetelményének meghatározására VaR alapú belső modelleket használjanak. Azok számára, akik nem akartak belső modellt fejleszteni, a szabályozó egyszerű algoritmusokat kínált a minimális tőkeszükséglet meghatározására. A sztenderd modellt az Egyesült Államok nem fogadta el, de javasolja azon kisebb intézményei számára, melyek európai piacokon vannak jelen. A sztenderd modell a tőkét különböző pozíciókhoz rendeli hozzá, amint azt az Európai Unió direktíva megfogalmazza, hogy fedezetet nyújtson ezen pozíciók kockázatára, így ez is egyfajta kockázati mértéknek tekinthető. Ezek nem rendelkeznek valószínűségi tartalommal, illetve nem is funkcionáljai a sűrűségfüggvénynek. Ráadásul ezen kockázati mértékek némelyike szintén megsérti a konvexitás követelményét (Kondor et al., [2004]). Ezen okok miatt nem végezhető el ezen mértékek és a variancia, illetve a többi kockázati mérték értelmes összehasonlítása a zajérzékenység tekintetében. Ugyanakkor a szabályozói mértékek rendelkeznek egy speciális tulajdonsággal, mely mind az Expected Shortfallt, mind a többi későbbiekben taglalt kockázati mértéket jellemzi: ezek szakaszonként lineáris mértékek, isorisk felületeik poliéderek.

## 12.2. Abszolút eltérés

Az abszolút eltérés mint kockázati mérték ( $E(|x|)$ , azaz az abszolút értékben vett eltérések várható értéke) egy kézen fekvő alternatívája a szórásnak. Néhány kockázatmenedzsment módszertan (pl. Algorithmics [2002]) ténylegesen az abszolút eltérést használja arra, hogy a portfóliók fluktuációját jellemezze, ami komoly számítási előnyöket jelent, mivel ez esetben az optimum egy lineáris programozási feladat megoldásaként áll elő. Ezzel szemben nehezen találunk irodalmat (Simaan [1997]) arra vonatkozóan, hogy milyen az átlagos eltérés zajérzékenysége.

Véve egy  $T$  hosszúságú véges idősort, a minimalizálandó célfüggvény (Konno et. al. [1991]) alapján a következő:

$$\frac{1}{T} \sum_t \left| \sum_i w_i x_{it} \right|, \quad (37)$$

az  $\sum_i w_i = 1$  feltételek mellett. (A hozamokra vonatkozó feltételektől ismét eltekintünk.)

Ez ekvivalens a következő lineáris programozási feladattal:

$$\min \frac{1}{T} \sum_t u_t \quad (38)$$

$$\text{s.t. } u^t + \sum_i w_i x_{it} \geq 0 \quad (39)$$

$$u^t - \sum_i w_i x_{it} \geq 0 \quad (40)$$

$$\sum_i w_i = 1 \quad (41)$$

(ahol a  $w_i$  és  $u_t$ ,  $t = 1 \dots T$  segéd változóknak optimalizálunk, míg az (39-40) egyenletek  $2T$  db korlátozó egyenletet jelentenek.)

Most alkalmazzuk ugyanazt a szimuláción alapuló stratégiát, mint a variancia esetén. Generáljunk mesterséges adatokat egy ismert struktúrából, az egyszerűség kedvéért ezek legyenek független standard normális eloszlású változók. Mivel standard normális

eloszlás esetén:  $E(|X|) = \sqrt{\frac{2}{\pi}} \sigma$ ,  $q_0$  AD esetén is a következőképpen számolható:

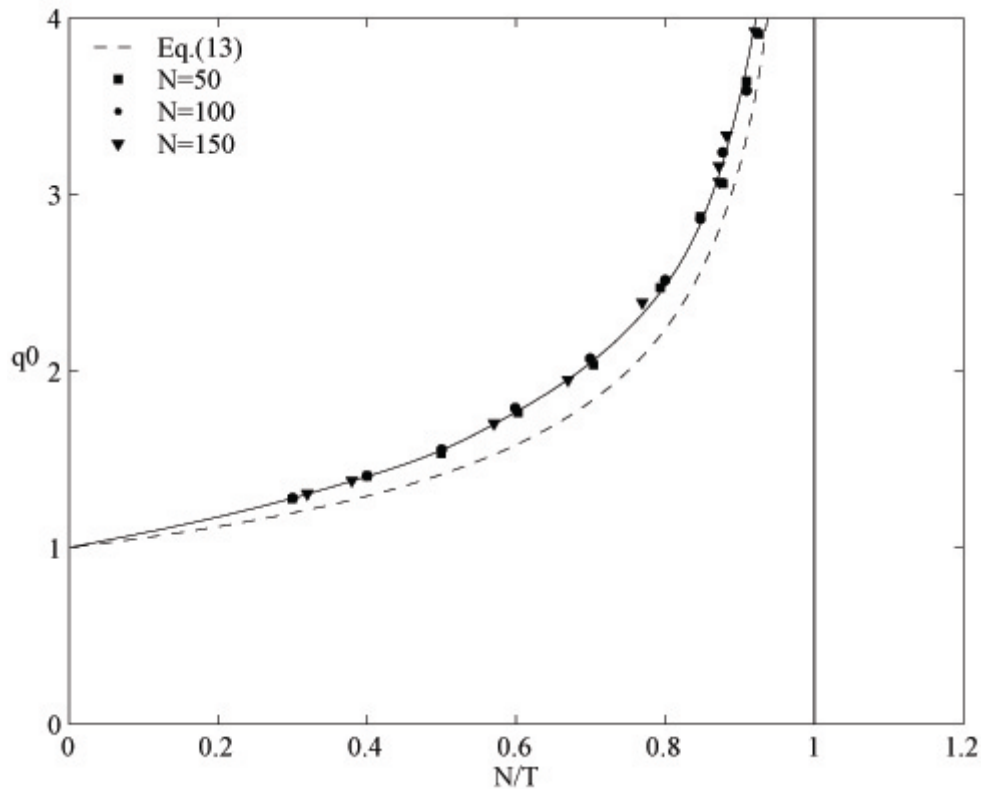
$$q_{0,AD}^2 = \frac{\sum_i w_i^{*2}}{\sum_i w_i^{(0)*2}}, \quad (42)$$

ahol a számláló a fenti optimalizációból származó súlyokon alapuló portfólió szórása, míg a nevező az igazi súlyokból álló portfólió szórása.

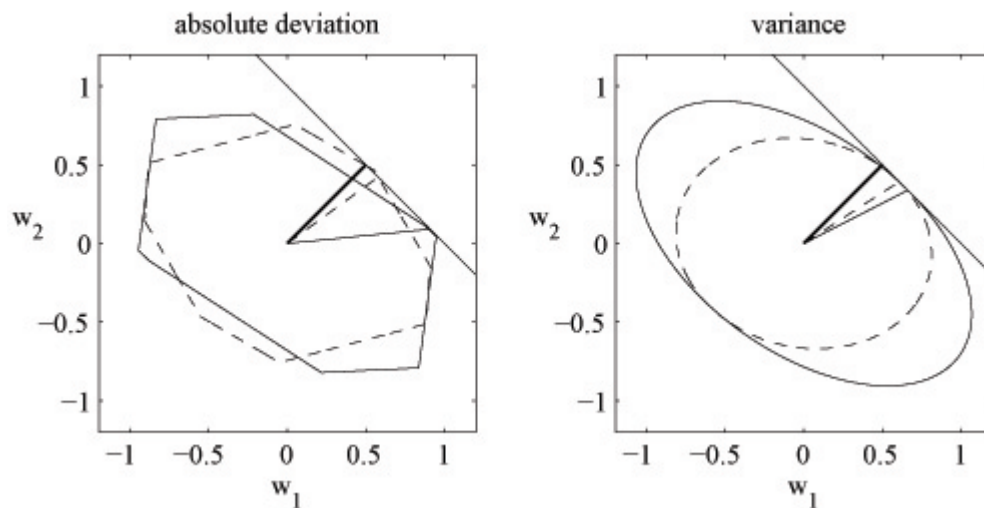
A szimmetria miatt az igazi súlyok mind egyenlőek ( $w_i^{(0)*} = \frac{1}{N}$ ), a  $q_0$  hányados pedig, ami a megfigyelt portfólió szuboptimalitását jellemzi, a következő formára egyszerűsödik:

$$q_{0,AD}^2 = N \sum_i w_i^{*2}. \quad (43)$$

A különböző  $N/T$  értékekre elvégzett szimuláció eredményét az 54. ábra tartalmazza. Látható, hogy az adatok újra egy görbére esnek, ami azt mutatja, hogy  $q_{0,AD}$  ismét  $N/T$ -ben skálázódik. Az ábra tartalmazza a variancia esetében megfigyelt eredményeket is. Egyértelműen kiderül, hogy  $q_{0,AD}$  a varianciát jellemző  $q_0$  fölött halad, ami azt jelenti, hogy az átlagos eltérés mint kockázati mérték érzékenyebb a zajra, mint a variancia. (Ez konzisztens Simaan [1997] eredményével.)



54. ábra:  $q_{0,AD}$  az  $N/T$  függvényében (szimulációs eredmények).  $q_{0,AD}$  pontjai is egy görbére esnek, mint azt a variancia esetében (szaggatott vonal) is tapasztalhattuk.



55. ábra: AD és a variancia iso-risk felületei  $N = 2$  és  $T = 4$  esetén. Az említett kockázati mértékek minimalizációs feladatainak a megoldását mutatja az ábra (vastag vonal jelzi az „igazi súlyokat”, míg a normál és szaggatott vonalak 1-1 optimalizáció eredményét mutatják).

A megnövekedett érzékenység okát Kondor et. al. [2006] megmutatta. Az isorisk felületek az átlagos eltérés esetén poliéderek, míg a variancia esetén ellipszoidok. Az optimalizáció megoldása az a pont, ahol ez a poliéder, vagy ellipszoid először érinti a korlátozó hipersíkot. Ez tipikusan a poliéder valamelyik extrémális pontjában történik meg. Ha véges idősből állítjuk elő ezt a poliédert, akkor az elkerülhetetlenül tartalmazni fog becslési hibákat. Következésképpen a poliéder alakja és/vagy a pozíciója mintáról mintára változni fog. Ennek eredményeképpen a megoldás egyik extrémális pontról a másikra fog ugrálni (lásd 55. ábra). Ez a diszkontinuitás az alapvető oka a megnőtt érzékenységnek. Amennyiben  $N/T$  tart a nullához, a poliéder ellipszoiddá alakul, és a különbség az abszolút eltérésben és a varianciában optimalizált portfólió között eltűnik. A mérték szakaszonként lineáris jellege igen vonzó tulajdonság, hiszen gyors és hatékony algoritmusok implementálását teszi lehetővé, ugyanakkor épp ez a kedvező jellemvonása okozza a nagyobb zajérzékenységet.

### 12.3. *Expected shortfall*

(önálló eredmény)

Az Expected Shortfall (lásd pl. Acerbi [2004]) egy valószínűségben definiált küszöbön túli várható érték. Abszolút folytonos eloszlások esetén ez ugyanaz, mint a feltételes VaR (Conditional VaR, CVaR), de diszkrét eloszlások esetén a két kockázati mérték a CVaR és az ES már nem ugyanaz. A különbségek részletes vizsgálatát többek között (Acerbi et. al. [2002]) mutatja be. Feltételes várható értéként az ES ésszerűbb jellemzését adja a nagy veszteségeknek, mint a VaR. Ezen kívül ez egy koherens kockázati mérték (ebből adódóan konvex is), nem úgy, mint a VaR, bár megjegyezzük, hogy azok közül a legegyszerűbb és a leginkább intuitív (Acerbi et. al. [2002]). Valójában a koherens kockázati mértékeknek egy speciális halmazába tartozik, melyet spektrális mértékeknek hívunk (Acerbi [2004]). Ezen vonzó tulajdonságok miatt az ES erősen támogatott azon személyek körében, akik aggodalmukat fejezik ki a VaR inkonzisztenciája miatt. Rockafellar et. al. [2000] megmutatták, hogy az ES optimalizációja lineáris programozási feladatra redukálható, mely igen hatékony algoritmusokat kínál.

Az ES célfüggvénye ( $v$ -ben és  $w_i$ -ben minimalizálva) (Rockafellar et. al. [2000]) alapján a következő:

$$\left( v + \frac{1}{(1-\beta)T} \sum_t \left[ -v - \sum_i \omega_i x_{it} \right]^+ \right), \quad (44)$$

$\sum_i w_i = 1$  feltételek mellett. (A hozamokra vonatkozó feltételeket elhagytuk, mint ezt az egész tanulmány során tesszük.) Az  $[\alpha]^+$  jelentése a következő: ha  $\alpha > 0$ , akkor értéke  $\alpha$ , egyébként 0, míg  $\beta$  az ES-hez tartozó konfidencia szint.

A fenti optimalizációs feladat ekvivalens a következő lineáris programozási feladattal:

$$\min \left( v + \frac{1}{(1-\beta)T} \sum_t u_t \right) \quad (45)$$

$$\text{s.t. } u_t \geq -v - \sum_i w_i x_{it} \quad (46)$$

$$u_t \geq 0 \quad (47)$$

$$\sum_i w_i = 1 \quad (48)$$

(ahol a  $w_i$ ,  $v$  és  $u_t$ ,  $t = 1 \dots T$  változóiban minimalizálunk).

Az ES zajérzékenységének tesztelésére független standard normális eloszlású változókat használtunk teljes analógiában a korábbi modellekkel. A véges  $T$  idősor fölötti optimalizálásból adódó szuboptimumot  $q_{0,ES}$ -sel jelöljük.

A szimuláció során egy nem várt jelenséggel találkoztunk: azzal, hogy a megoldás nem mindig létezik! Konkrétabban a megoldás létezése a mintától függ, és ezáltal egy valószínűségszámítási kérdéssé válik. A megoldás létezésének valószínűsége az  $N$ ,  $T$  és  $\beta$  paraméterek függvénye. Mielőtt belevágnánk az ES zajérzékenységének mérésébe, szükségesnek érezzük tisztázni ezt a problémát, és kiválasztani azokat a területeket, melyeken azt várjuk, hogy a megoldás közel 1 valószínűséggel létezik. Mivel a feladat igen bonyolult, valamint hogy előzetes betekintést nyerhessünk a probléma



$\beta$  1-hez közel van. Ez

esetben  $(1-\beta)T \leq 1$ , ami azt jelenti, hogy csak a legkisebb elem járul hozzá az ES-hez. Ez a leszűkített eset így önmagában is egy koherens (spektrális) mértéket ad, a maximális veszteséget. A megoldhatósági probléma ML esetén is jelen van, és az ES-sel ellentétben analitikusan könnyen követhető, azaz egyféle laboratóriumként szolgál a jelenség modellezésére.

#### 12.4. Maximális veszteség

(önálló eredmény)

A maximális veszteség mint kockázati mérték talán túl pesszimistának tűnhet, lényege, hogy egy adott  $T$  idősor fölött a maximális veszteséget minimalizáljuk a súlyokban. A minimalizálandó célfüggvény a következő:

$$\max_t \left( \sum_i w_i x_{it} \right), \quad (49)$$

a szokásos súlyokra vonatkozó korlátozás mellett  $\sum_i w_i = 1$ .

A minimax probléma (Young [1998]) ekvivalens a következő lineáris programozási feladattal:

$$\min u \quad (50)$$

$$\text{s.t. } u \geq - \sum_i w_i x_{it} \quad (51)$$

$$\sum_i w_i = 1 \quad (52)$$

A portfólió szuboptimalitását  $q_{0,ML}$ -lel jelöljük. Jelentése analóg a korábban definiáltakkal. Mint azt már említettük, az optimum nem mindig létezik, ezért  $q_{0,ML}$ -t azon feltételezés mellett kell értenünk, hogy a feladat optimalizálható.

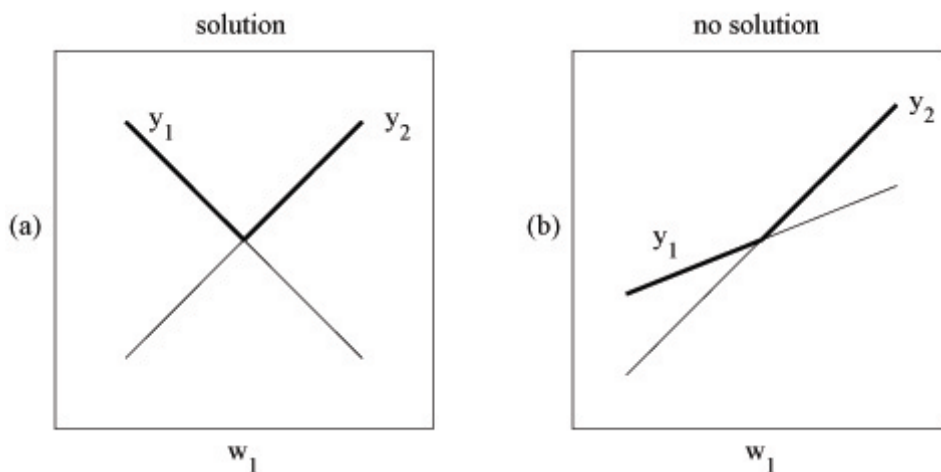
Ezt a megoldhatósági problémát a következő egyszerű példán keresztül kívánjuk megvilágítani: képzeljük el a következő egyszerű esetet, ahol  $N = 2$ , és  $T = 2$ , ekkor két lineáris függvényünk van a 26-os egyenletből adódóan:

$$y_1 = -w_1 x_{11} - (1 - w_1) x_{21} = (x_{21} - x_{11}) w_1 - x_{21} \quad (53)$$

$$y_2 = -w_1 x_{12} - (1 - w_1) x_{22} = (x_{22} - x_{12}) w_1 - x_{22}, \quad (54)$$

ahol a  $w_2$  súlyt a korlátozó egyenletbe történő behelyettesítéssel már elimináltuk. Ha a két fenti egyenesnek a meredeksége ellenkező előjelű, azaz  $x_{21} > x_{11}$  és  $x_{22} < x_{12}$ , vagy  $x_{21} < x_{11}$  és  $x_{22} > x_{12}$ , akkor létezni fog a megoldás, más esetben azonban nem. Az 56. ábra mutatja a fent leírt eseteket. Könnyen látható, hogy független standard normális eloszlású hozamok esetén a megoldás létezésének valószínűsége  $1/2$ .

Most már értjük a probléma gyökerét: az ML kialakításánál fogva egy konvex kockázati mérték, de a problémát az okozza, hogy alulról nem mindig korlátos, és ez esetben nem létezik az optimum. Ugyanakkor az is világos, hogy ha korlátokat vetünk ki a súlyokra, azaz megtiltjuk, hogy negatívak legyenek, ez megszünteti a problémát (ez a short selling megtiltását jelenti).



56. ábra:  $y_1(w_1)$  és  $y_2(w_1)$ -nek a szövegben leírt esetei. (a): az ML kockázati mérték alulról korlátos, és a megoldás létezik, (b): a mérték nem korlátos, így a minimumhely nem létezik.

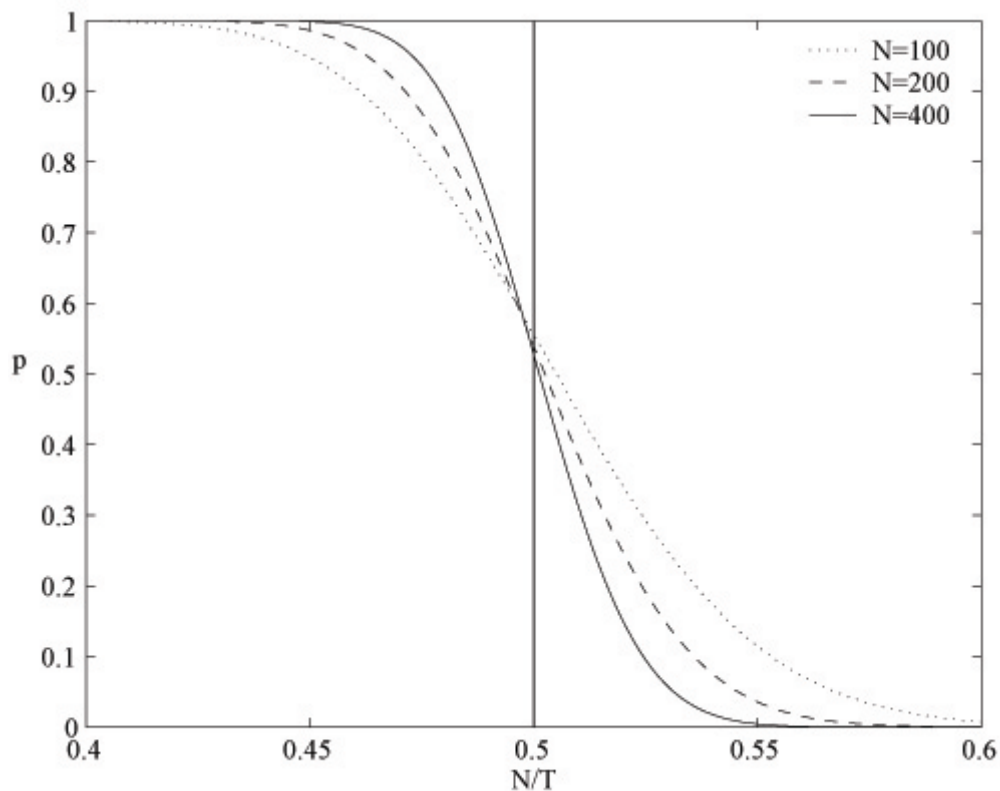
Általános esetet vizsgálva  $T$  darab véletlen hipersíkunk van egy  $N$  dimenziós térben, és a megoldás létezésének feltétele, hogy ezek a hipersíkok egy konvex politópot alkossanak.  $N$  és  $T$  figyelembe vételével az optimum létezésének a valószínűségére a következő érvényes:

$$p = \frac{1}{2^{(T-1)}} \sum_{k=N-1}^{T-1} \binom{T-1}{k}. \quad (55)$$

A korábbiakban független standard normális eloszlású változókkal dolgoztunk, az eredmények érvényesek tetszőleges többváltozós normális eloszlásra, sőt elliptikus eloszlásokra is. A kockázati mérték optimalizálhatóságának valószínűsége nem függhet a koordináta rendszer tengelyeinek az irányától, sem pedig skálázástól, ami pl. a független normális eloszlást elliptikusba viszi át.

Hasonló probléma már korábban is felmerült a sztochasztikus programozás és a véletlen geometria területén (lásd Todd [1991]; Mattheiss et. al. [1977]).

Az 57. ábra illusztrálja az optimalizálhatóság valószínűségét  $N/T$  függvényében,  $N$  különböző értékeire.  $N$ -et növelve az átmenet egyre élesebbé válik, míg a limeszben ( $N; T \rightarrow \infty$ ) az átmenet megszakad, és  $p$  egy lépcsős függvénnyé válik.



57. ábra: Annak a valószínűsége, hogy a minimax problémának létezik megoldása  $N/T$  függvényében  $N$  különböző értékei mellett. Ahogy  $N \rightarrow \infty$ ,  $p$  lépcsős függvénné válik.

$N$  és  $T$  nagy értékeire a valószínűség eloszlás jól közelíthető a következő függvénnyel:

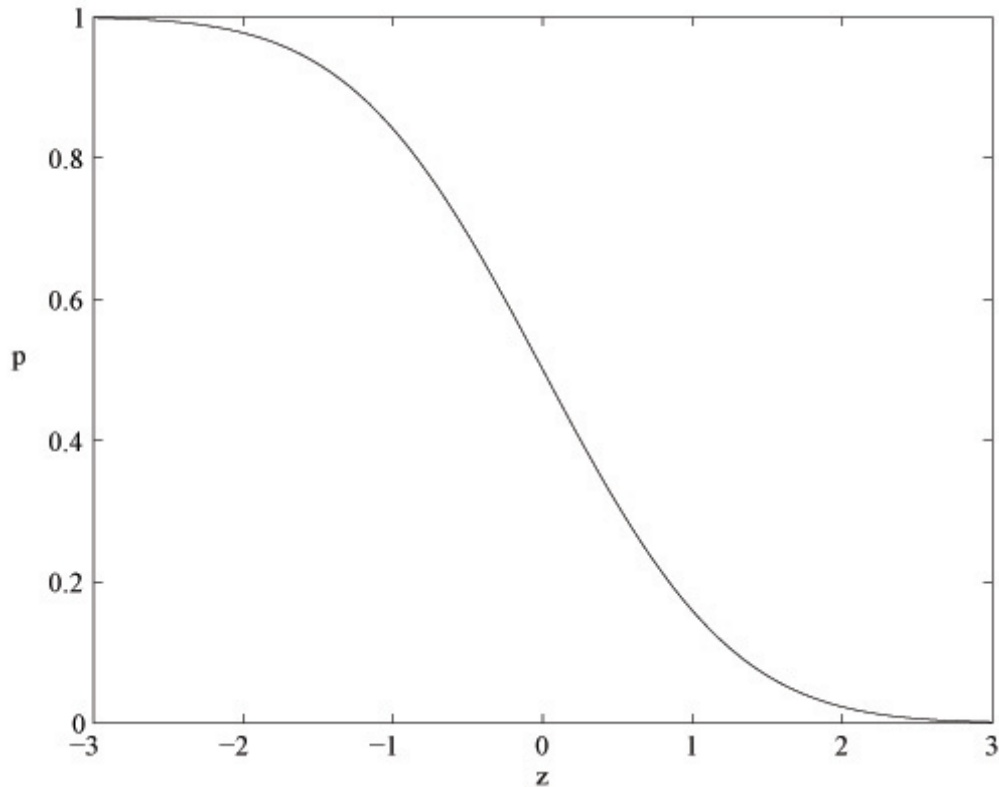
$$P \approx 1 - \Phi(z), \text{ ahol} \quad (56)$$

$$\Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}s^2} ds, s \quad (57)$$

$$z = 2 \left( \frac{N}{T} - \frac{1}{2} \right) \sqrt{T}. \quad (58)$$

Az 58. ábraán a megoldhatóság valószínűségét ábrázoljuk  $z = 2 \left( \frac{N}{T} - \frac{1}{2} \right) \sqrt{T}$  függvényében. Ha  $N$  és  $T$  egyaránt nagyok, és  $N < T/2$ , akkor  $z$  nagy és negatív, és a megoldás létezésének valószínűsége 1-hez nagyon közel van. Amikor  $N > T/2$ , akkor  $z$  nagy és pozitív, és  $p$  0-hoz közele. Az átmenet egyre élesebbé válik  $T$  növekedésével. Az  $N/T = 1/2$  speciális esetet képvisel, hasonlóan az  $N/T = 1$ -hez a variancia esetében. Mivel

a kovariancia mátrix maximális rangja  $N$  és  $T$  közül a kisebbik, az optimalizáció a varianciában mindig

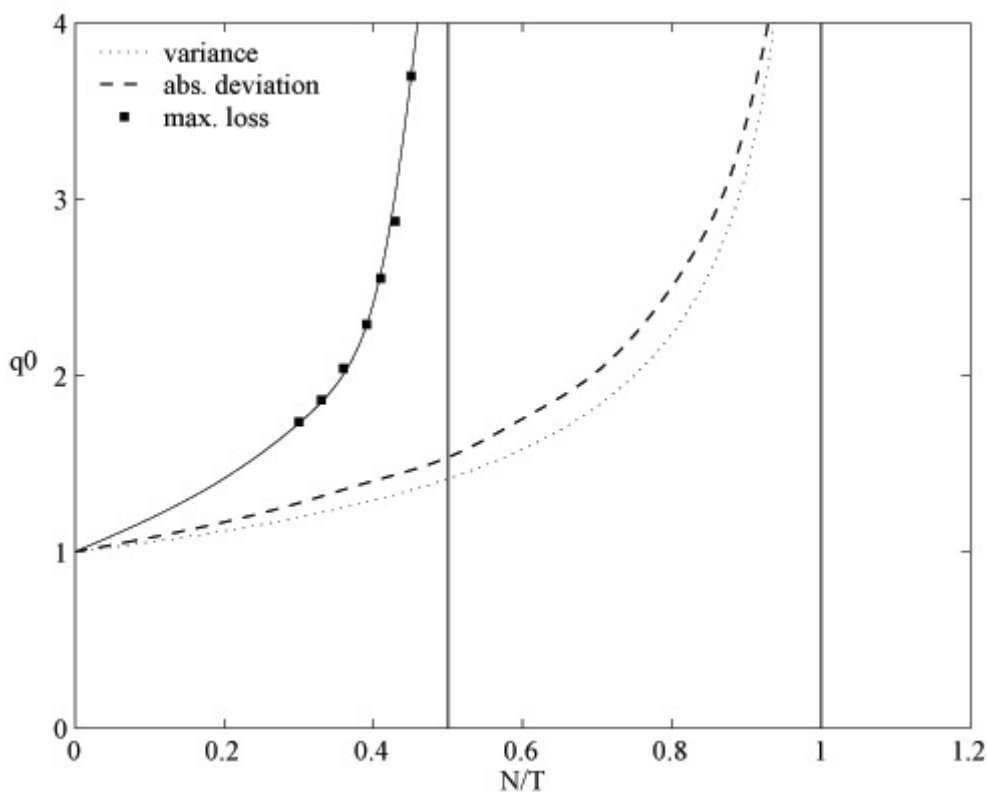


58. ábra: Annak a valószínűsége, hogy a minimax feladatnak létezik megoldása  $z = 2\left(\frac{N}{T} - \frac{1}{2}\right)\sqrt{T}$  függvényében.

lehetséges, ha  $N/T < 1$ , és mindig értelmetlen, ha  $N/T > 1$ . A megoldás létezésének valószínűségét bevezethetjük tehát a variancia esetén is, ami egy lépcsős függvény lesz, és a függvény  $N/T = 1$  esetén ugrik. A különbség ML esetében az, hogy véges  $N/T$ -re az átmenet folytonos, ami azt jelenti, hogy sosem lehetünk teljesen biztosak abban, hogy a megoldás létezik. A konvergencia ugyanakkor az eloszlás mindkét végén nagyon gyors.

Feltéve, hogy elegendő mennyiségű olyan adat áll rendelkezésünkre, melyeken a megoldás létezik, futtathatjuk a korábban bevezetett szimulációkat, és megmérhetjük  $q_{0,ML}$ -t,  $N/T$  függvényében. Egyáltalán nem kell meglepődnünk, ha azt tapasztaljuk, hogy  $q_{0,ML}$  sokkal nagyobb, mint  $q_0$  a variancia esetében, vagy  $q_{0,AD}$ .

ML isorisk felületei is poliéderek, mint azok AD esetében voltak, és a megoldás itt is egyik extrémális pontról a másikra ugrál zajos környezetben. Sokkal fontosabb azonban, hogy ML-ben történő optimalizálás esetén egy napi (időegységnyi) adat kivételével az összes többit eldobjuk. Ezek után nem csoda, hogy a becslés nagyon instabillá válik ilyen körülmények között.



59. ábra:  $q_{0,ML}$   $N/T$  függvényében (szimulációval kapott eredmények). Az összehasonlíthatóság kedvéért az ábra a variancia és az AD korábban bemutatott  $q_0$  értékeit is tartalmazza.

Az 59. ábra ML zajérzékenységét mutatja, a varianciához és az AD-hez viszonyítva. Látható, hogy a három közül ML a legérzékenyebb, és az  $N/T = 1/2$  kritikus értéknél  $q_{0,ML}$  divergál.

### 12.5. Az expected shortfall megoldhatósági problémája

(önálló eredmény)

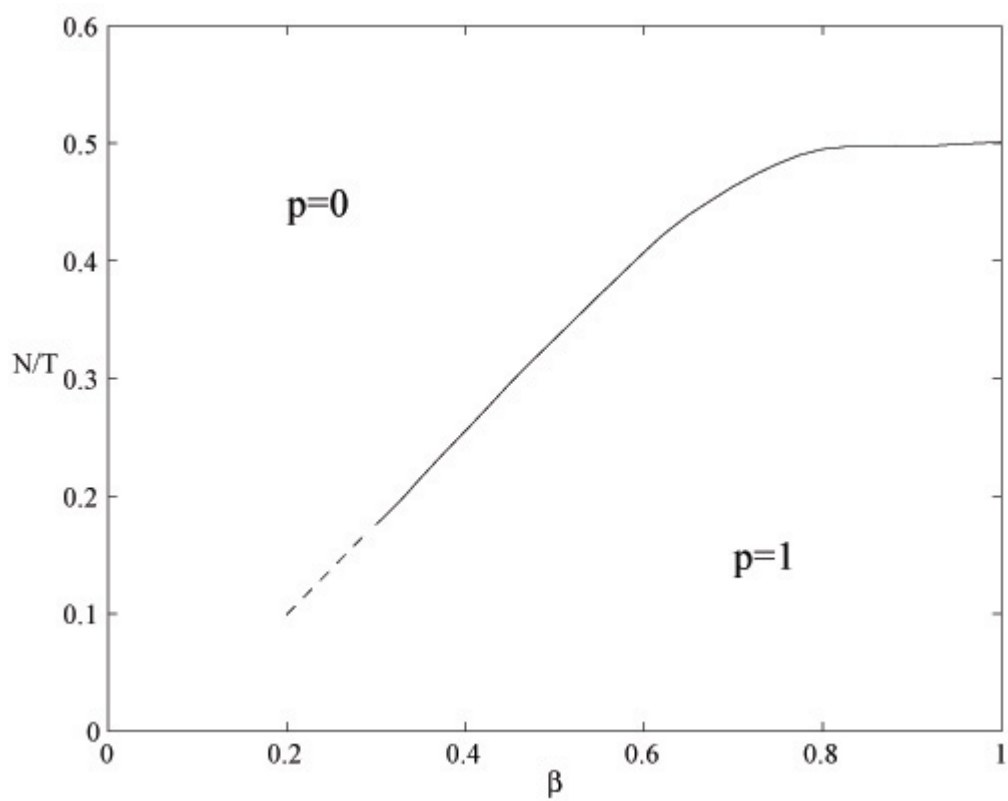
Most már képesek vagyunk az ES megoldhatósági térképének a fölrajzolására. Kísérleteink eredménye a „fázisdiagram”, melyet a 60. ábra tartalmaz. A

megoldhatósági térkép annak az esetnek felel meg, amikor  $N$  és  $T$  a végtelenbe tartanak, miközben arányuk folyamatosan változik 0 és 1 között a függőleges tengely mentén. A konfidenciaszint ( $\beta$ ) változását a vízszintes tengelyen tüntettük fel. Az előző alfejezetből tudjuk, hogy  $\beta = 1$  esetén (ami az ML) a megoldás létezésének a valószínűsége 1-ről 0-ra esik le  $N/T = 1/2$ -nél, ahogy haladunk a  $\beta = 1$  egyenes mentén. Ahogy eltávolodunk  $\beta = 1$ -től,  $N/T$  kritikus értéke (azaz ahol a megoldás valószínűsége ugrik) csökken. Ez azt jelenti, hogy rögzített  $N$  mellett egyre hosszabb és hosszabb idősort kell figyelembe vennünk ahhoz, hogy a feladat ugyanakkora eséllyel megoldható legyen. Ez azt is jelenti, hogy ES egyre érzékenyebbé válik a zajra ahogy  $\beta$  csökken. Mivel  $\beta$  csökkenése azzal jár, hogy egyre több és több adatot veszünk figyelembe, ez az eredmény igen meglepő.

A fázisátmenet határának pontos alakját bonyolult meghatározni, különösen  $N/T$  kis értékeire, ahol a szimulációk már igen lassúkká válnak. Ezen ok miatt a fázisátmenet függvényt már nem mértük ki  $\beta = 0,3$ -nál kisebb értékekre, viszont könnyű megmutatni, hogy ennek az origóban kell végződnie, mivel  $\beta > 0$  és  $N/T \rightarrow 0$  esetén minden információnk rendelkezésre áll, emiatt szükségszerűen  $\rho = 1$ , míg  $N/T > 0$  és  $\beta \rightarrow 0$  esetén a minimalizálandó célfüggvény lineárisává válik, és emiatt sosem lesz korlátos a súlyokban, következésképpen az optimalizálhatóság valószínűsége 0.  $\beta$  kis értékeinek nincs gyakorlati jelentősége, mi csupán azt szeretnénk kihangsúlyozni, hogy a fázisátmenet határa lefelé kanyarodik, ahogy  $\beta$  értéke egytől a nullához tart.

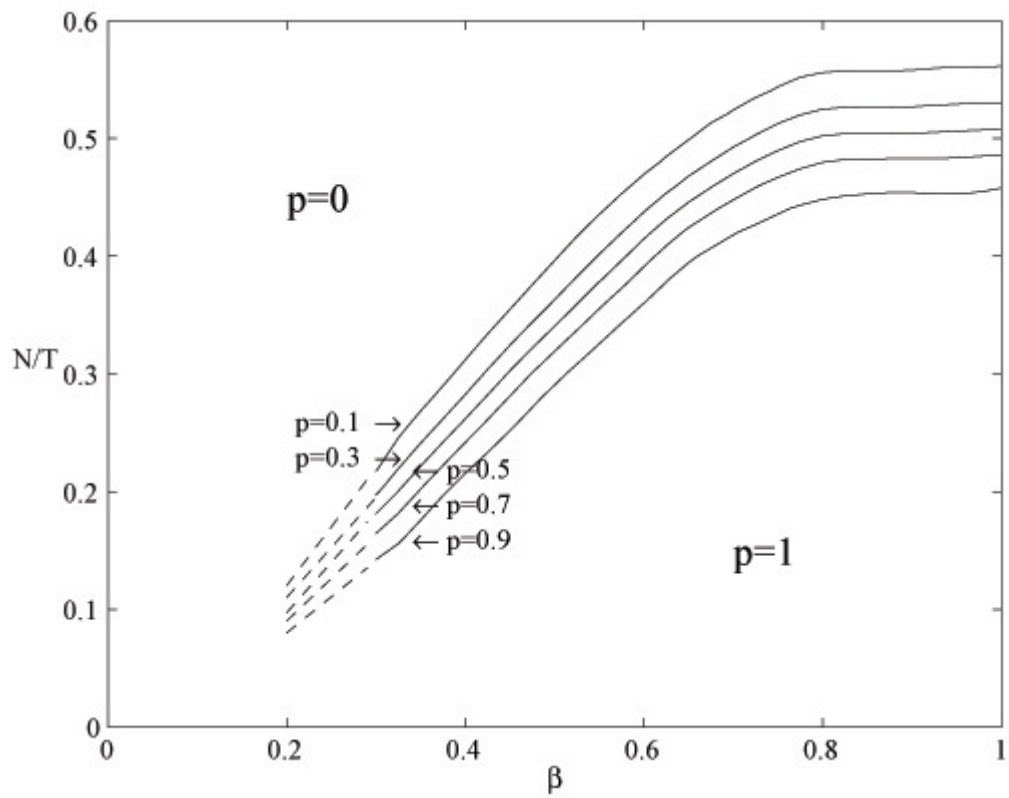
Véges  $N$ ,  $T$  esetén a függvény éles zuhanása, ami a limeszben ( $N, T \rightarrow \infty$ ) volt megfigyelhető, folytonos átmenetté alakult, hasonlóan az ML esetén megfigyeltékhez. Ezt illusztrálja a 61. ábra, ahol az iso-p görbék ( $p(\beta, N/T)$ ) láthatók  $N = 100$  esetén, míg a 62. ábra  $p$  változását mutatja  $N/T$  függvényében, rögzített  $\beta = 0,7$  mellett. A short selling megtiltásával a megoldhatósági probléma itt is megszűnik.

Ezek után már vizsgálhatjuk az ES zajra való érzékenységét. Rögzített  $\beta$  mentén haladva (feltéve, hogy a megoldás létezik)  $N/T$  növekedése mellett  $q_{0,ES}$  is gyorsan nő, és divergálni kezd a kritikus  $N/T$  érték közeledtével (lásd 63. ábra).



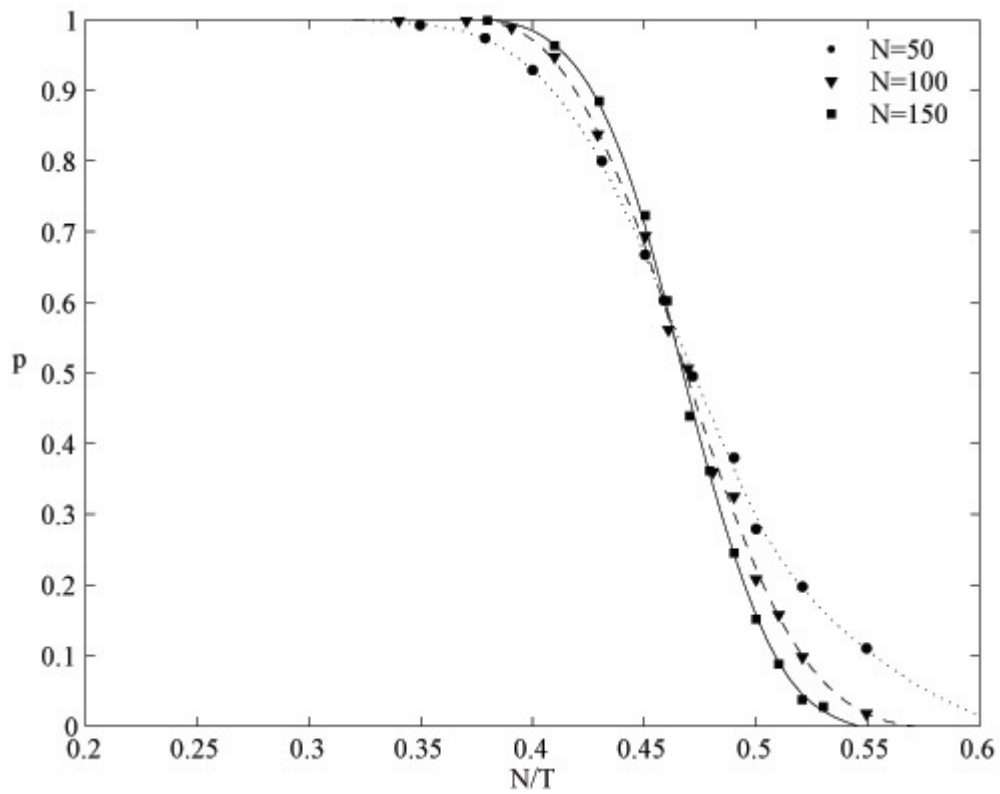
60. ábra: A fázisdiagramm a  $p = 0$  és  $p = 1$  állapotokat, és a köztük lévő határvonalat mutatja a  $\beta$  - $N/T$  térben, miközben  $(N, T \rightarrow \infty)$ .



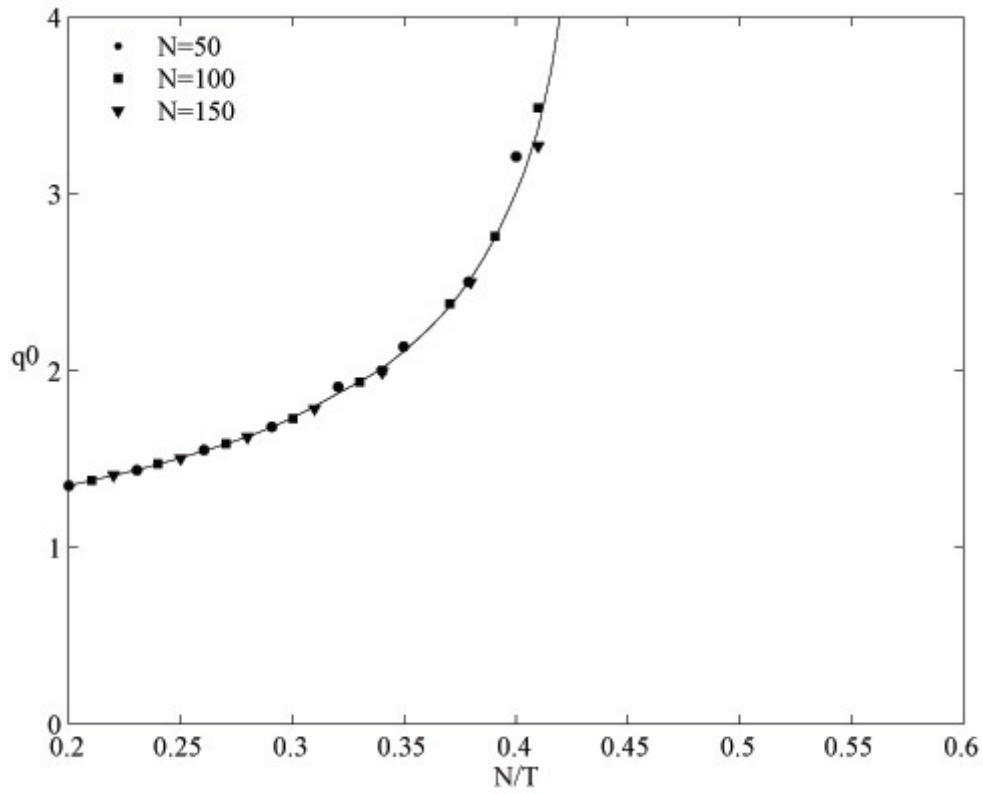


61. ábra:  $p$  szintvonalai  $\beta$  és  $N/T$  függvényében  $N = 100$  esetén

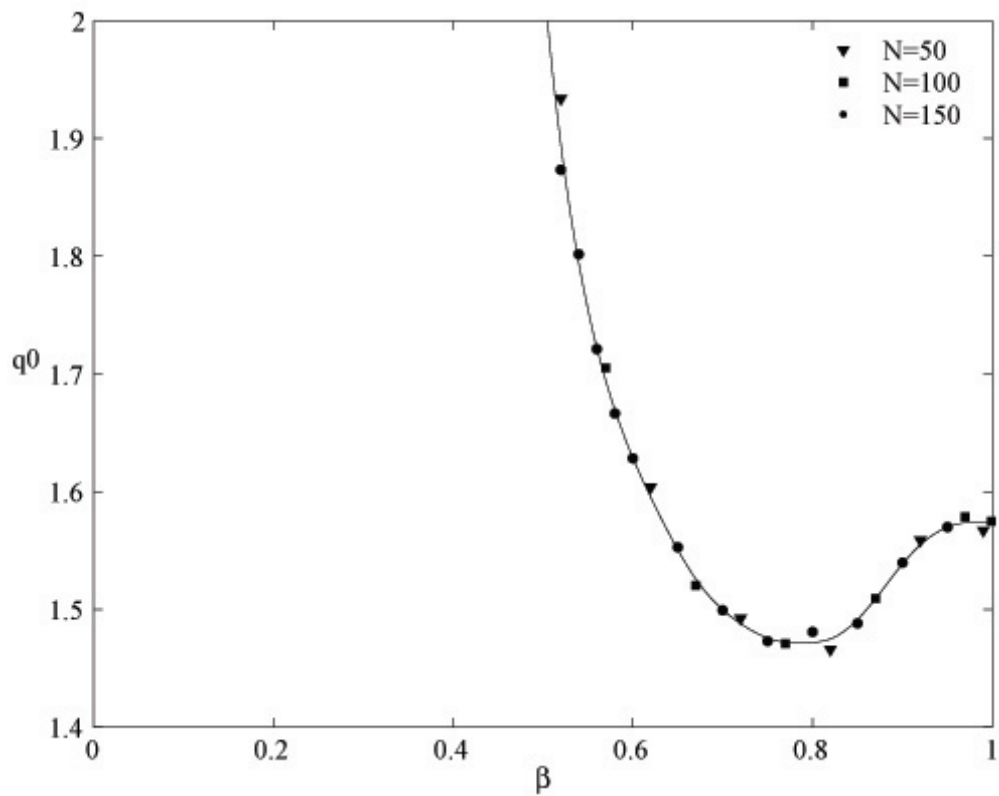
Ezen túlmenően még egy érdekes tényvel találkozhatunk, ha rögzített  $N/T$  esetén  $q_{0,ES}$  változását vizsgáljuk  $\beta$  függvényében. (Ez megfelel annak, hogy a



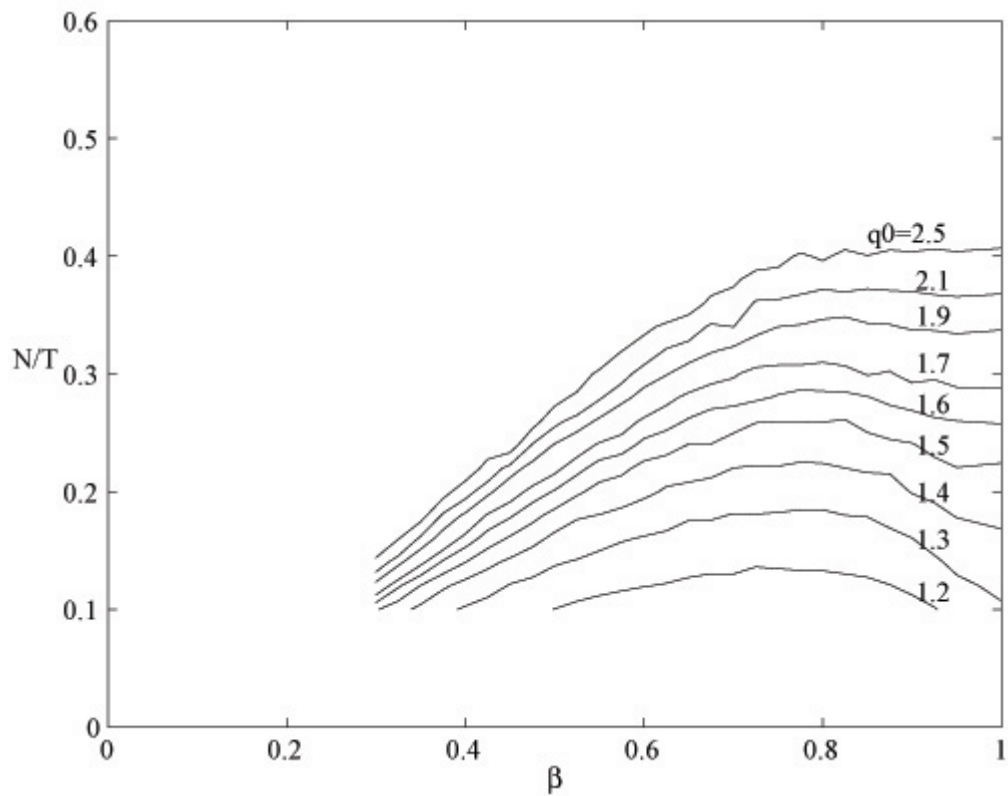
62. ábra: Annak a valószínűsége ( $\rho$ ), hogy az ES feladat megoldható  $N/T$  függvényében,  $\beta = 0:7$  értéke mellett ( $N$  különböző értékeire).  $N \rightarrow \infty$  esetén  $\rho$  lépcsős függvénné válik.



63. ábra:  $q_{0,ES}$   $N/T$  függvényében  $\beta = 0.7$  rögzített értéke mellett (különböző  $N$ -ekre).

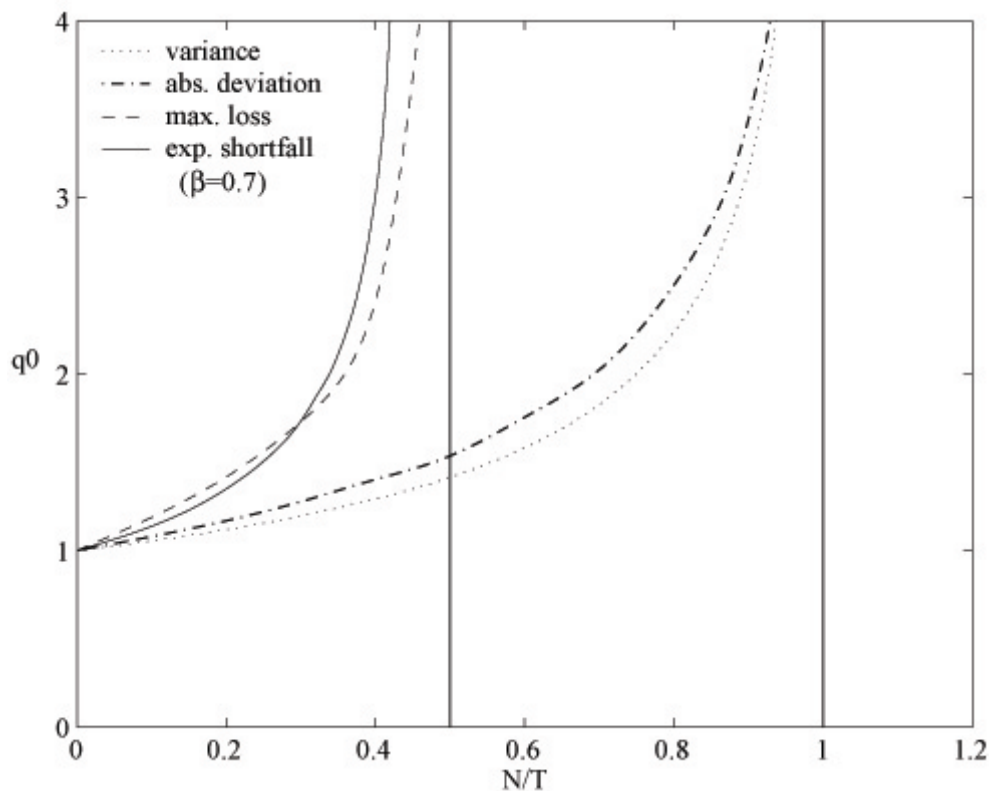


64. ábra:  $q_{0,ES}$   $\beta$  függvényében rögzített  $N/T = 1/4$  mellett ( $N$  különböző értékeire).



65. ábra:  $q_0(\beta, N/T)$  szintvonalai

60. ábraán a vízszintes tengely mentén haladunk.) Ahogy  $\beta$  értéke csökken, kezdetben  $q_{0,ES}$  is vele együtt csökken, de miután elérte a minimumát, újból növekedni kezd, és a fázisátmenet határán divergál (lásd 64. ábra).  $q_{0,ES}$ -nek ez a nem monoton változása  $\beta$  függvényében a 65. ábraán látható, ahol  $q_0(\beta, N/T)$  szintvonalai találhatók meg.



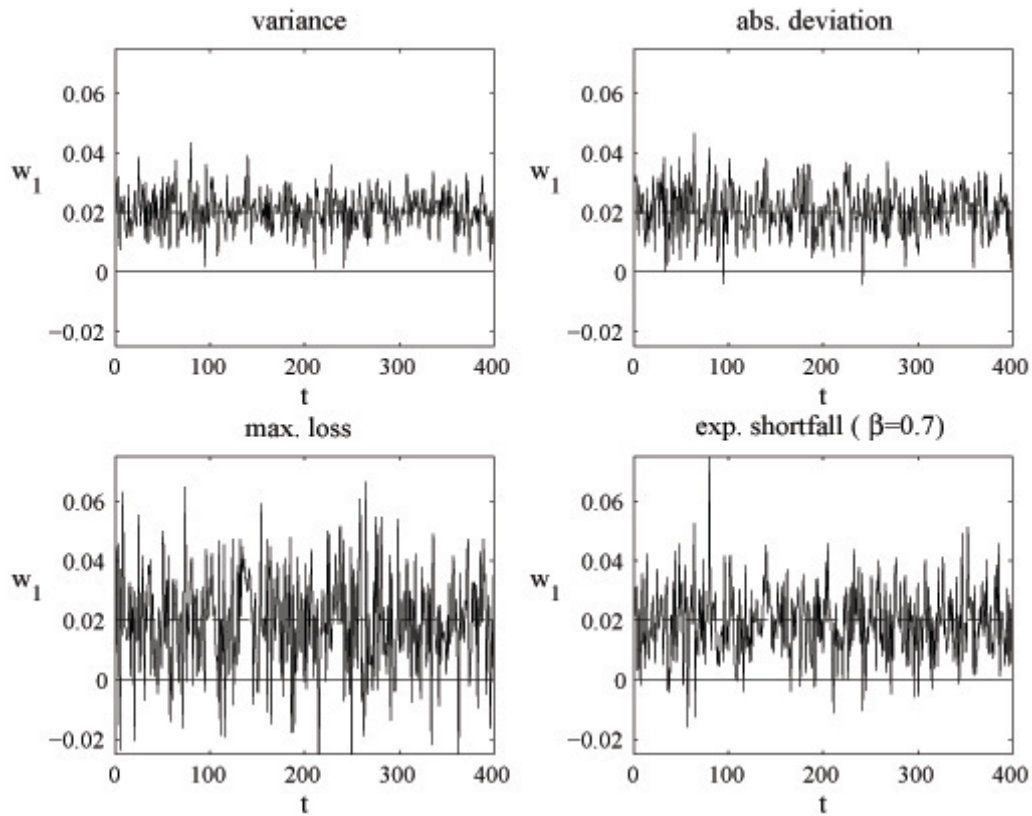
22. ábra.  $q_0$   $N/T$  függvényében az ismertett kockázati mértékek esetén

Mind  $q_{0,ES}$ , mind pedig a fázisátmenet határok viselkedését érdekesnek találjuk, mivel azt vártuk, hogy az ES egyre kevésbé és kevésbé lesz érzékeny a zajra  $\beta$  csökkenésével, ami azt jelenti, hogy egyre több és több adatot veszünk figyelembe. Amikor a különböző kockázati mértékek zajérzékenységét vizsgáltuk, világos volt számunkra, hogy nem fair a varianciát egy magas konfidenciaszintű ES-sel összehasonlítani, mivel ezáltal a megfigyelt adatok nagy részét eldobjuk. (Köszönettel tartozunk ezen észrevételért Carlo Acerbinek.) Ami meglepő, hogy  $\beta$  csökkenésével a megoldhatóság valószínűsége csökken,  $q_0$  pedig a kezdeti csökkenés után növekedésnek indul. Tekintettel  $q_{0,ES}$  ezen nem monoton viselkedésére, a legkedvezőbb összehasonlítást akkor tudnánk megtenni, ha  $\beta$  és  $N/T$  terében kijelölnénk azt a részt, amely mentén  $q_{0,ES}$  minimális. Sajnálatos módon ez óriási számítási munkát igényelne, emiatt erre nem került sor. A rendelkezésre álló eredmények alapján azonban sejthető, hogy ES egy megnövekedett zajérzékenységet mutatna nem csak a varianciával, hanem AD-vel szemben is, még a legkedvezőbb körülmények esetén is.

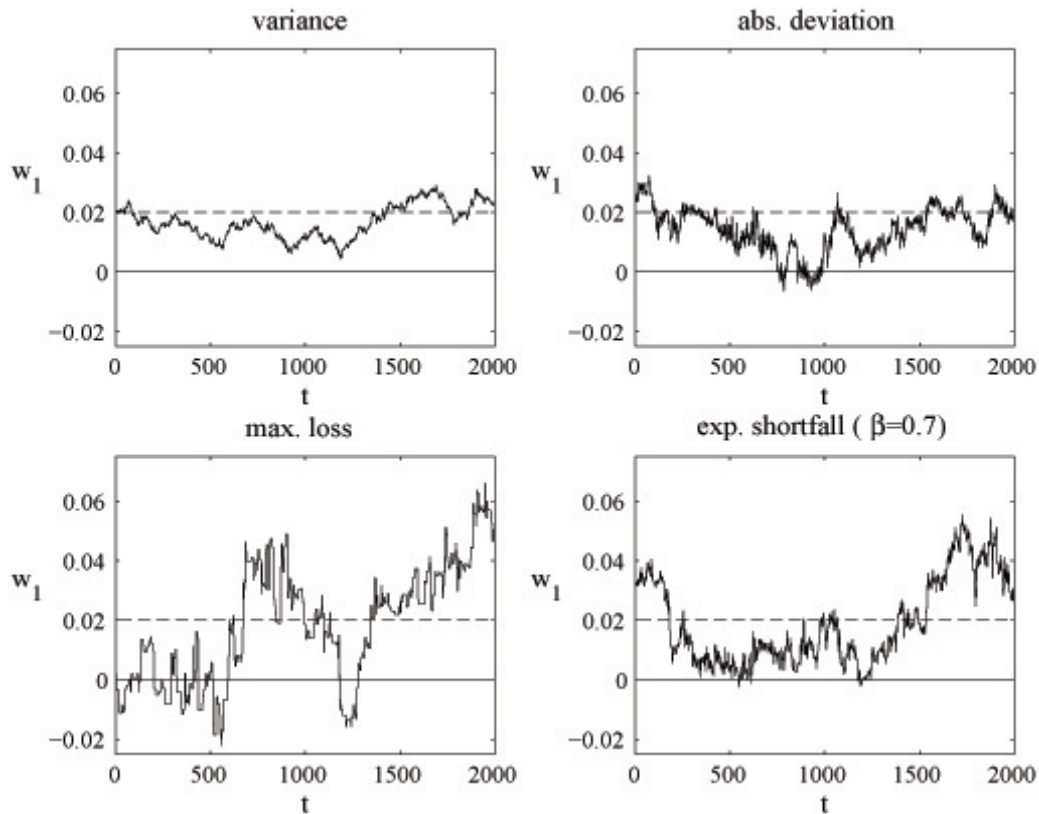
Az ebben a fejezetben leírt eredmények mind független standard normális eloszlású változók felhasználásával születtek. Könnyű belátni, hogy az állítások korrelált normális

## 12.6. A súlyok ingadozása és $q_0$ eloszlása

Vessünk még egy pillantást a portfóliók zaj okozta instabilitására, és tanulmányozzuk a súlyok ingadozását, illetve  $q_0$  eloszlását. A korábbiakban vázolt szimulációkat újra végrehajtva a 66. ábra és 67. ábra eredményeihez jutottunk (mind a négy kockázati mérték esetén korrelálatlan normális eloszlású változókat használva.)



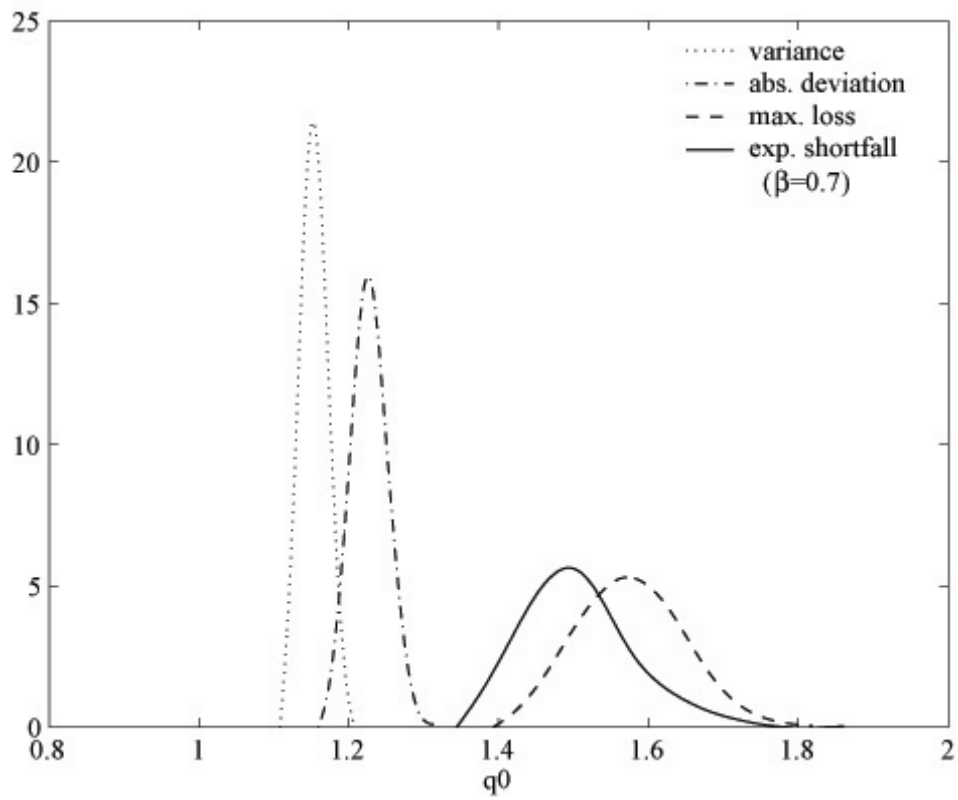
66. ábra:  $\omega_1^*$  az idő függvényében mind a 4 kockázati mérték esetén, egymást nem átfedő idősorokból számítva ( $N = 50$ ,  $T = 500$ ).



67. ábra:  $\omega_1^*$  az idő függvényében mind a 4 kockázati mérték esetén, egymást átfedő idősorokból számítva, azaz az idősort 1 egységgel folyamatosan léptetve ( $N = 50$ ,  $T = 500$ ).

Amint azt az ábrán láthatjuk, még a variancia esetén is igen instabilak a súlyok, még akkor is, amikor maga  $q_0$  elfogadható lenne. Nem átfedő idősorok esetén a mintáról mintára történő változások igen nagyok és korrelálatlanok, míg egymást átfedő (egy időponttal csúsztatott) idősorok esetén a súlyok erősen korreláltak, és egy zajtól perturbált útvonalat járnak be, míg az igazi súlyok konstansok. Az is látható, hogy a súlyok instabilitása egyre inkább növekszik, ahogy az egyik kockázati mértékről a másikra váltunk, a következő sorrendben: variancia, AD, ES, ML.

Ugyanezt a következtetést vonhatjuk le a 68. ábrát szemlélve is, ami azt mutatja, hogy a  $q_0$ -ok eloszlása egyre kiterjedtebbé válik, ahogy a varianciától a maximális veszteség felé tartunk.



68. ábra:  $q_0$  eloszlása  $N = 200$ ,  $N/T = 1/4$  esetén a tanulmányban kifejtett kockázati mértékekre (szimulációs eredmények).



### *13. Konklúziók a portfólióoptimalizálásban használt kockázati mértékek instabilitásáról*

A tipikus banki portfólió eszközök nagy száma, és a limitált adatmennyiség következtében a zaj a portfólióelméletben átható nehézséget jelent. Az optimalizálás során fellépő hibával foglalkozó tanulmányok többsége a varianciára, mint kockázati mértékre, valamint a valós környezetre, az empirikus adatokra fókuszál. A dolgozatban a portfóliók zajérzékenységét különféle kockázati mértékek alatt optimalizálva vizsgáltuk, és hasonlítottuk össze. Kockázati mértékként kezeltük a varianciát, de az abszolút szórást (absolute deviation), az Expected Shortfallt és a Maximális Veszteséget (Maximal Loss) is. A portfólió érzékenységére globális mérőszámot alkalmaztunk, és a különböző kockázati mértékeket szimulált portfóliókon teszteltük, ahol portfóliók méretét (N) és az idősorok hosszát (T) változtattuk.

Az eredmények azt mutatják, hogy a zaj hatása az összes vizsgált esetben jelentős volt, és nagy mértékben a kockázati mértékenként különböző N/T arányon múlt. Mind a fenti négy kockázati mértékre meghatároztuk ezt a kritikus értéket, az Expected Shortfall is beleértve, ahol N és T minden határon túl történő növelése esetén egy egyértelmű határ alakult ki, elkülönítve azt a részt, ahol a portfólióoptimalizálás 1 valószínűséggel következik be, attól a résztől, ahol a megoldás megtalálásának valószínűsége nulla.

A zajérzékenységnek egy másik aspektusa a mintáról mintára fluktuáló portfóliósúlyok. Nagy fluktuációkat találtunk már variancia mellett optimalizált portfólióknál is, de a többi kockázati mérték esetén ezen fluktuációk még mindig jelentősebbnek bizonyultak. A súlyok instabilitása látható, ha egymást nem átfedő mintákat használunk. Az egymást átfedő minták autokorrelációkat generálnak, itt az optimum lassan változik, ugyanakkor valójában egy szuboptimális helyzetben reked meg a portfólió.

Az elemzés egyik fő mondanivalója, hogy a variancián kívüli kockázati mértékek esetén több információ szükséges: ugyanazon portfólióméretre, és a hozamok ugyanazon (normál) eloszlása esetén több adatra, pl. hosszabb idősorokra van szükség, hogy a varianciával versenyképesek legyenek. Nyilvánvaló, hogy normál eloszlású hozam esetén a variancia a legjobb kockázati mérték. Mivel az empirikus adatok azonban nem normál, és időben változó eloszlásúak, azt gondoljuk, hogy a becslési hiba hatásai a valóságból vett portfóliók esetén csak rosszabbak lehetnek, mint a mesterséges adatok esetén.

Míg a variancia esetén számos hatékony szűrési módszert fejlesztettek ki a zaj hatásának kiküszöbölésére, az alternatív kockázati mértékek esetében alig van néhány szűrő módszer. A más kockázati mértékek szerint optimalizált portfóliók megnövekedett zajérzékenységének megfigyelésekor láthatjuk, hogy egyértelműen szüksége lenne olyan szűrő módszerekre, melyek az adott kockázati mértékre adaptáltak.

A dolgozatban használt kockázati mértékek a variancia kivételével szakaszonként lineárisak. A linearitás nagy számítási előnyöket nyújt, de láttuk, hogy ugyanakkor instabilitáshoz is vezet: a minta változása az eredmények „ugrálásához” vezet, ami a portfóliósúlyok megnövekedett fluktuációjában manifesztálódik.

Az Expected Shortfall és annak extrém esetének, a maximális veszteségnek az esetén különösen erős a mintáról mintára történő fluktuáció, és azzal a meglepő megvalósíthatósági problémával találkoztunk, hogy bár az ES egy konvex kockázati mérték, bizonyos minta fölött az optimalizációnak nincs megoldása. Megjegyezzük, hogy Kondor et. al. [2010] általánosságban is megmutatta ezt a tulajdonságot a koherens kockázati mértékekre. A megvalósíthatósági problémát ki lehet küszöbölni további feltételek alkalmazásával, például short selling kizárásával. A megoldás valószínűsége a probléma paraméterein múlik; maximális veszteségnél ez analitikusan megmutatható, míg az Expected Shortfall esetén kimértük azt határvonalat, ahol,  $N$ ,  $T$  tart a végtelenhez esetén a megoldás 1 valószínűséggel létezik. Meglepő eredményként értékelhető az ES esetében, hogy  $q_0$  nem monoton módon viselkedik  $\beta$  függvényében.

Habár matematikai szempontból számos érdekes tulajdonságot mutattunk be, a befektető szemszögéből nézve a legfontosabb üzenetünk, hogy a különböző kockázati mértékek nem egyformán érzékenyek az optimalizálás során elkövetett hibára, ami az idősorok véges jellegéből adódik. Napjainkban a kockázati mérték választása során igen divatos szempont, hogy olyan mértéket válasszunk, melyek tipikusan a nagy veszteségekre érzékenyek, azaz a kockázat fogalma a nagy veszteségekhez kötődik. Ugyanakkor ezen mértékek (esetünkben az ES és ML) mellett a portfólió optimalizálása csak nagy hibával végezhető el. Ennek következménye, hogy míg a befektető azt gondolja, hogy egy adott hozamszint mellett a minimális kockázatú portfóliót választotta, addig a választott portfólió kockázata többszöröse lehet a befektető által véltnek. Az optimális portfólió választása során alkalmazott modellek esetében kockázati mértéket tehát nem csupán abból a szempontból kell választanunk, hogy ahhoz milyen kockázati fogalom társítható, hanem a mérték és portfóliósúlyok becslése során fellépő hiba mértéke is döntő szempont kell, hogy legyen. A konklúziókat Paul Embrechts kérdésével zárjuk, amikor a

40: „Nem lehet, hogy mégiscsak a szórás a legjobb kockázati mérték?”

#### *14. Saját eredmények összegzése*

Az alábbiakban összefoglaló jelleggel felvázoljuk a dolgozat főbb kutatási kérdéseit. Mivel az azokra adott válaszokat a dolgozat konklúzió fejezeteiben már összegeztük, így itt elsődlegesen azon eredményekre térünk ki, melyek a szerző önálló eredményének tekinthetők.

Fontosnak tartjuk azonban megemlíteni, hogy a dolgozat eredményei nem csupán a problémafelvetésekre adott válaszokban tükröződnek, hiszen számos egyéb esetben is (az elemzés melléktermékeként) adódtak új, vagy újszerű, illetve a nemzetközi szakirodalommal összhangban lévő eredmények. A kérdésfelvetéseket és az azokhoz legszorosabban kapcsolódó eredményeket félkövérrel kiemelve közöljük.

A dolgozat három fő kutatási kérdésfelvetés köré épül:

#### **I. A működési kockázatok tökekövetelményéhez kapcsolódó extrapolációs probléma ugyanúgy érinti-e a kis és nagy mintával rendelkező intézményeket?**

Ezt a problémát három megközelítésben vizsgálatuk.

- i. Építettünk egy modellt a fenti probléma minél egyszerűbb kezelésére. Ebben a modellben vizsgálatuk a kis és nagy minták tulajdonságait.
- ii. Elemeztük az empirikus adatokat a HUNOR működési kockázati adatbázis vizsgálatán keresztül. Az adatbázis anonim, így eredeti célunk az volt, hogy a külső adatok hatását vizsgáljuk a kis és nagy intézményekre.
- iii. Mivel a HUNOR adatbázis számos empirikus problémát támasztott, így mesterséges, szimulációs környezetben vizsgáltuk a külső adatok jelentőségét a kis és nagy mintával rendelkező intézmények tőkeszámítására.

A modellépítés (i) konklúziói az alábbiak szerint összegezhetők.

---

<sup>40</sup> Paul Embrechts egy budapesti látogatása során beszélgettünk az előzetes eredményekről a Collegium Budapest épületében. A beszélgetésen még Kondor Imre és Pafka Szilárd vettek részt.

**1. Tézis.** *Lognormális egyedi veszteségeloszlás mellett a tőkekövetelmény szintén lognormális eloszlású (felételezve, hogy ismerjük a gyakoriság eloszlás várható értékét, valamint az MLE becslés klasszikus tulajdonságai teljesülnek - részletesen lásd 79. oldal).*

Ez megerősíti azt az empirikus tényt, miszerint a tőkekövetelmény számításban elkövetett hiba nem szimmetrikus, hanem jobbra ferde. A levezetés jelentősége az empiria megerősítésén túl abban rejlik, hogy a szimulációs módszernél lényegesen gyorsabb utat kínál az eloszlás megismerésére (az adott modellfeltételek mellett).

**2. Tézis.** *Csupán belső adatokat figyelembe véve, az extrapolációból adódó hátrányokat kompenzálja a növekvő mintaelemszám (adott modellfeltételek mellett).*

A bemutatott modell környezetben rávilágítottunk, hogy a becslés hibája (VaR szórása) kevésbé vastag szélű ( $\sigma$  kicsi) egyedi káreloszlás esetén monoton csökkenő a gyakoriság függvényében, míg növekvő  $\sigma$  mellett a kezdeti csökkenés megtorpan és  $\lambda$  (gyakoriság várható értéke) magasabb értékei mellett újra emelkedni kezd. Mindez érdekes helyzetet teremt a tőke becslése kapcsán. Nem csak az számít, hogy mennyi adattal rendelkezik egy intézmény, hanem az is, hogy mennyire vastag szélű a veszteséggeneráló folyamat. A kettő összjátéka határozza meg, hogy mennyire pontosan tudja az intézmény megbecsülni a működési kockázati tőkekövetelményét.

Mint azt már korábban említettük két hatás feszül egymásnak. Egyrészt a növekvő gyakoriság kedvez a becslésnek, hiszen nagyobb mintából a veszteség eloszlást kisebb hibával tudjuk becsülni, másrészt viszont a nagyobb gyakoriság egyre távolabbi kvantilis extrapolálását teszi szükségessé. Úgy tűnik, hogy míg egy kevésbé kockázatos folyamat (kis  $\sigma$ ) esetén a mintanagyság növekedés ellensúlyozni tudja az extrapolációból adódó bizonytalanságot, addig heavy tail folyamatok esetében ez már nem teljesül. Ebből adódóan paradox módon a mintanagyság növekedése a tőke becslésének nagyobb bizonytalanságával párosul heavy tail folyamatok esetében!

Természetesen a kép ennél árnyaltabb, hiszen a növekvő szórás még mindig lehet elhanyagolható a tőkéhez képest, így érdemes a relatív szórásra is egy pillantást

vetni. A relatív szórás vizsgálata kapcsán annak monoton csökkenő jellegét tapasztaltuk. Ez azt jelenti, hogy a több veszteség eseménnyel rendelkező bankok relatíve mégiscsak jobb helyzetben vannak - eloszlástól függetlenül, hiszen a tőke nagyságrendjéhez viszonyított hiba monoton csökken. Tehát az extrapolációból adódó hátrányokat kompenzálja a növekvő mintaelemszám (az adott modellfeltételek mellett)!

Lognormális egyedi veszteségeloszlású modellünkben (bizonyos feltételezések mellett) megmutattuk, hogy a VaR eloszlása is lognormális, és várható értéke és szórása egyaránt függ a mintaelemszámtól ( $n$ ). A kevés veszteséggel rendelkező intézmények várhatóan sokkal drasztikusabban becslik túl a működési kockázati tőkéjüket, mint a sok veszteséggel rendelkezők. Ez is azt a logikát erősíti meg, hogy az extrapolációs probléma hatása nem egyforma az egyes intézmények esetében. Minél kevesebb adattal rendelkezik egy intézmény, várhatóan annál inkább túlbecsli a működési kockázati tőkéjét.

Elemzésnek vetettük alá a HUNOR működési kockázatkezelési adatbázist (ii), ahol a korábbi hazai elemzésektől eltérő kontextusban vizsgáltuk azt. Az ebből adódó eredmények elsődlegesen a nemzetközi szakirodalommal mutattak összhangot, kutatási kérdésünk megválaszolásában nem vezetett előre.

Az elemzés rávilágít, hogy az egyes eseménykategóriák veszteségei vastag szélű eloszlást követnek. Ezen túlmenően a lognormális eloszlás a széleken rosszul illeszkedik, és tipikusan alulbecsli a kockázatokat. Rávilágítottunk, hogy az egyes paraméter becslési módszerek ugyanazon eloszlás feltételezése mellett is akár egy nagyságrenddel eltérő tőkekövetelményeket adhatnak!

Az empirikus adatok vizsgálata (általánosított Pareto eloszlás illesztése) megmutatta, hogy az MLE becslőfüggvény igen érzékeny az azonos mintaelemekre, azok kezelése mindenképp szükséges.

Szimulációs környezetben vizsgáltuk a külső adatok jelentőségét a kis és nagy intézmények tőkeszámítására (iii), köztük hogy az általánosított Pareto eloszlás (külső adatok eloszlása) paramétereit, az intézmény veszteségeinek átlagos száma, külső adatbázis mérete, becslési módszerek (MLE, PWM) hogyan hatnak az intézmény tőkekövetelményére. Ezzel kapcsolatban számos, a nemzetközi

irodalommal (empirikus megfigyelésekkel, egyéb modellekkel) összhangban lévő megállapításokat tettük, valamint kutatási kérdéseinkre egy újabb aspektusból (belső adatok kiegészítése külső adatokkal) adunk választ.

**3. Tézis.** *A külső adatok figyelembe vétele (feltéve, hogy mindegyik intézménynek ugyanaz a külső minta áll rendelkezésére) nagyobb előnnyel jár (relatív nagyobb mértékben tudják csökkenteni a tőkekövetelmény becslési hibáját) a kisebb intézmények számára, mint a nagyobbak számára.*

A nagyobb gyakorisággal rendelkező intézmény nem feltétlenül van jobb helyzetben a kisebbekhez képest. Az éves gyakoriság növelésével a VaR relatív szórása (szórás / VaR hányados) nőtt mind az MLE, mind a PWM esetében,  $\xi$  minden értékére! Ugyanakkor a relatív szórás mértéke az MLE becslés esetében csaknem egy nagyságrenddel kisebb, nagy  $\xi$  esetén. A várható értékkel normált tartományok, melybe a VaR 98%-os valószínűséggel esik, mind az MLE mind a PWM becslés esetén nőttek nagyobb gyakoriság mellett! Felhívjuk a figyelmet, hogy az extrapolációs problémának most egyértelműen a negatív hatásával szembesül egy sok belső adattal rendelkező intézmény egy kevés adattal rendelkezőhöz képest. A nagyobb belső mintaelemszám miatt távolabbra kell extrapolálnia az eloszlás szélére a VaR megállapításához, mint egy kevesebb belső adattal rendelkező intézménynek, miközben mindkettőjüknek ugyanakkora minta (külső adatbázis) áll rendelkezésre az eloszlás szélén! Ez pedig rontja a becslés jóságát.

Szimulációs vizsgálatunk számos olyan eredménnyel járt, melyek ugyan nem járulnak hozzá közvetlenül az első számú kérdésfelvetés megválaszolásához, ugyanakkor számos, a nemzetközi szakirodalommal (empirikus megfigyelésekkel, egyéb modellekkel) összhangban lévő megállapítást tartalmaznak. Ezeket az alábbiakban összegezzük.

- Tapasztaltuk az MLE becslés (aszimptotikus) torzítatlanságát, míg a PWM becslőfüggvény várhatóan alulbecsli az igazi tail index paramétert. A torzítás mértéke annál nagyobb (PWM esetén) minél közelebb van az igazi paraméter (tail index) 1-hez.

- A tail index MLE becslése szimmetrikus, míg a PWM jobbra ferde.
- A PWM tail index becslése felülről korlátos, míg az MLE nem.
- Minél közelebb van  $\xi$  (tail index) értéke az 1 -hez annál inkább különbözik az MLE és a PWM becslés.
- A tail index MLE becslése esetén nagyjából 5% annak a valószínűsége, hogy a becslés során egy nem véges várható értékű eloszláshoz jutunk ( $\xi=0,9$  esetén).
- A scale paramétert a PWM becselőfüggvény felülbecsli, míg az MLE (aszimptotikusan) torzítatlan becslést ad.
- A scale paraméterre vonatkozó 90%-os konfidencia sáv a PWM esetében szélesebb.
- Méréselt a különbség az egyes becselőfüggvényekhez tartozó scale paraméter eloszlások között, ezt erősíti meg a paraméter becslési hibájának a tőkére gyakorolt hatása is.
- A nagyobb  $\beta$  (scale paraméter) nagyobb szórással is jár, de a relatív szórások már megegyeznek.
- A PWM becselőfüggvény a tőkét várhatóan alulbecsli, míg az MLE becslés aszimptotikusan torzítatlan.
- A 90%-os biztonsággal becsült tőke alsó és felső értéke között megközelítőleg 1 nagyságrendnyi eltérés van, még mérsékelt tail index mellett is!
- A 98%-os biztonsággal becsült tőke minimuma és maximuma között 30-szoros eltérés is előfordult! A maximum és minimum hányadosa minden esetben kisebb az MLE becslés esetén, mint a PWM módszer használatával. Szembetűnő azonban az a jelenség, hogy  $\xi=0,9$  esetén a PWM becsléshez tartozó VaR eloszlásának felső (0,99) kvantilise kisebb, mint az MLE becsléshez tartozó VaR ugyanazon kvantilise. Ennek oka a  $\xi$  eloszlásának vizsgálatokor megfigyelt jelenség lehet, hogy a PWM becslés során  $\xi$

- Minél közelebb van  $\xi$  értéke 1-hez, annál inkább különböznek az egyes becslőfüggvények által adott VaR eloszlások.
- A tőkekövetelmény relatív szórása nő, ahogy  $\xi$ -vel 1-hez közelítünk. PWM esetében érdemes megfigyelni, hogy a szórás a várható érték többszörösére nő  $\xi=0,9$  esetén!

Röviden, az alábbiak szerint foglalhatjuk össze az első (I) kutatási kérdésre adott válaszunkat. Egy általunk épített modellezési környezetben megmutattuk, hogy az intézmény belső mintáját tekintve előnyt élveznek a nagyobb mintával rendelkező intézmények, hiszen a VaR relatív hibája csökken a minta elemszámának növelésével. Ezzel szemben a külső adatbázis mérete fix (ugyanannyi külső adat áll minden szereplő rendelkezésére), és mivel a kevesebb adattal rendelkező intézmények esetében kisebb az extrapolációs hatás, így annak negatív következményeit jobban tompítja egy ugyanolyan méretű külső minta, mint egy sok adattal rendelkező intézményét, ahol még a külső adatok jelentős száma is kevés lehet a stabil becsléshez.

Összességében még az általunk felállított, és kellően leegyszerűsített modellezési környezetben belül sem tudunk egy nyilvánvaló és egyértelmű választ adni a kutatási kérdésre. Mindez azt jelenti, hogy a gyakorlatban az egyes intézményeknek a körülményeket (külső-belső mintanagyság, veszteséggeneráló folyamat tulajdonságai, stb.) jól körbejárva, megértve és számos egyedi döntést meghozva kell működési kockázati tőkemodelljüket felépíteniük. A problémára adott aktuáriusi módszerek (LDA) ugyan gyors és kényelmes utat ígértek, ugyanakkor egyre nyilvánvalóbbnak tűnik, hogy nincs egyértelmű válasz a fenti problémára.

**II. A portfólióválasztási feladatok megoldását mennyiben befolyásolja a választott kockázati mérték? Avagy az egyes mértékek mennyire érzékenyek az optimalizálás során elkövetett becslési hibára?**



**4. Tézis.** *Az egyes kockázati mértékek becslései különbözőképpen érzékenyek a zajra. A vizsgált mértékek közül az Expected Shortfall és annak speciális esete, a Maximális Veszteség esetében a legnagyobb a becslési hiba, meghaladva az Absolute Deviation és a variancia esetén tapasztaltakat.*

Mivel a szórásra és abszolút eltérésre vonatkozó eredmények a szakirodalomban már rendelkezésre álltak, így elsődlegesen az Expected Shortfallt és annak speciális esetét a Maximális Veszteséget vizsgáltuk. Láttuk, hogy a vizsgált kockázati mértékek különbözőképpen érzékenyek a zajra. Azt tapasztaltuk, hogy az Expected Shortfall és annak speciális esete, a Maximális Veszteség esetében a legnagyobb a becslési hiba, meghaladva az AD és a variancia esetén tapasztaltakat. A portfóliósúlyok nagy fluktuációját láthatjuk már variancia mellett optimalizált portfólióknál is, de a többi vizsgált kockázati mérték (AD, ES, ML) esetén ezen fluktuációk még jelentősebbnek bizonyultak.

A súlyok instabilitása is jól látható, ha egymást nem átfedő mintákat használunk. Az egymást átfedő minták autokorreláltak, így itt az optimum lassan változik, ugyanakkor valójában egy szuboptimális helyzetben reked meg a portfólió.

Az elemzés egyik fő mondanivalója, hogy a variancián kívüli kockázati mértékek esetén több információ szükséges: ugyanazon portfólióméretre, és a hozamok ugyanazon (normál) eloszlása esetén több adatra, pl. hosszabb idősorokra van szükség, hogy a varianciával versenyképesek legyenek. A nagyobb információigény egyik oka a mértékek (AD, ES, ML) szakaszonként lineáris volta. A linearitás nagy számítási előnyöket nyújt, de láttuk, hogy ugyanakkor instabilitáshoz is vezet: a minta változása az eredmények „ugrálásához” vezet, ami a portfóliósúlyok megnövekedett fluktuációjában manifesztálódik. Ezen túlmenően az ES és ML mértékek a rendelkezésre álló idősor csupán egy részét, az extrémeket használják fel. Ebből adódik, hogy a kisebb mintából nagyobb becslési hiba mellett lehet az optimalizálást elvégezni.

Meglepő eredményként értékelhető az ES esetében, hogy  $q_0$ , ami az optimalizálás során elkövetett hibát méri nem monoton módon viselkedik  $\beta$  (valószínűségben definiált küszöb) függvényében.

**III. Létezik-e minden esetben optimális portfólió? Azaz a különböző portfólióoptimalizálási feladatoknak létezik-e minden esetben megoldása?**

**5. Tézis.** *Az Expected Shortfall és a Maximális Veszteség csupán adott valószínűséggel optimalizálhatók.*

Az Expected Shortfall és annak extrém esetének, a Maximális Veszteségnek az esetén különösen erős a mintáról mintára történő fluktuáció mellett azzal a meglepő megvalósíthatósági problémával találkoztunk, hogy bár az ES egy konvex kockázati mérték, bizonyos minta fölött az optimalizációnak nincs megoldása. Ezt a megvalósíthatósági problémát ki lehet küszöbölni további feltételek alkalmazásával, például short selling kizárásával, de ez már az eredetitől eltérő lineáris programozási feladatot jelent.

A megoldás valószínűsége a probléma paraméterein múlik; maximális veszteség ez analitikusan megmutatható, míg az Expected Shortfall esetén kimértük azt határvonalat, ahol,  $N$ ,  $T$  tart a végtelenhez esetén, a megoldás 1 valószínűséggel létezik.

## *15. Befejezés*

A dolgozat során azt a kérdést feszegettük, hogy napjaink kockázati mértékei a portfólióoptimalizálás során mennyire állják meg a helyüket, illetve a működési kockázati tőkekövetelmény meghatározása milyen nehézségekbe ütközik? Közös pont a két kérdésben, hogy mindkét esetben nagyfokú instabilitást rejtenek a modellek, annak alkalmazója közel sem lehet biztos a modell által adott eredmények jóságában. Az instabilitás forrása különböző a két esetben. Portfólióoptimalizálás esetén tudnunk kell, hogy a szakaszonként lineáris kockázati mértékek gyors optimalizálást tesznek lehetővé, ugyanakkor az isorisk felületek poligonok, ahol a megoldás az extrémális pontokban létezik. Ennek következtében az input adatokban bekövetkező, akár kismértékű változás a megoldás egyik extrémális pontból a másik extrémális pontba történő ugrásával járhat. Ez egyfajta zaj növekedést jelent a szakaszonként lineáris kockázati mértékek esetében. Az Expected Shortfall és a Maximális Veszteség esetében figyelembe kell vennünk továbbá azt is, hogy az optimalizálás során csak az input adatok egy részét vesszük figyelembe, hiszen célunk épp egy extrém veszteségeket minimalizáló portfólió összeállítása. Az információk nagy részének az eldobása további

zajnövekedéshez vezet. Ehhez járul hozzá az a jelenség, hogy ezen két kockázati mérték optimalizálása csak adott valószínűséggel végezhető el, a portfólió méretének, az idősor hosszának és az alkalmazott mérték, vagy a mérték paramétereinek a függvényében.

Működési kockázatok tekintetében az instabilitás forrása maga az eredeti Basel II-es szabályozói definíció, hiszen az extrapolációs problémával minden intézmény szembesül, működési adatbázisának méretétől függetlenül. A dolgozat során világossá vált, hogy nincs egyértelmű válasz arra a kérdésre, hogy a kis, vagy nagy mintamérettel rendelkező intézmények tudnak-e jobb becslést adni a tőkekövetelményükre, mivel ez számos tényező függvénye.

Összességében arra szeretném felhívni az olvasó figyelmét, hogy a pénzügyi modellek naiv, a probléma természetének megértése nélküli alkalmazása azt eredményezheti, hogy tévhitben élünk a dolgok valós kockázatának tekintetében, hiszen a valós kockázat akár többszöröse is lehet a modellek által kimutatottnak.

*Dolgozatban felhasznált saját publikációk jegyzéke*

Gáll, J., Nagy, G., 2007: [A működési kockázat veszteségeloszlás-alapú modellezése \(Loss Distribution Approach, LDA\)](#)

Hitelintézeti Szemle, 2007/4. szám

Hajnal, B., Kállai, Z., Nagy, G., 2007: [Működési kockázati önértékelések veszteségeloszlás-alapú modellezése](#). Hitelintézeti Szemle. 2007/5. szám

Kondor, I., Pafka, Sz., Karádi, R., Nagy, G., 2006. Portfolio selection in a noisy environment using absolute deviation as a risk measure. In: Takayasu, H. (Ed.), Practical Fruits of Econophysics. Springer.

Kondor, I., Pafka, Sz., Nagy, G., 2007: Noise sensitivity of portfolio selection under various risk measures, Journal of Banking and Finance, 31, 1545-1573 (2007).

Nagy, G., Povilaitis, K., 2009: [A működési kockázatok veszteségeloszlás alapú modellezésének lehetőségei](#). Hitelintézeti Szemle. 2009/2. Szám

## *Summary*

During the essay we have tried to find out how well today's risk measures hold their grounds during portfolio optimization, and what problems in the determination of operational risk capital requirements must be overcome? In both cases models contain considerable instability, their users cannot be sure at all about the quality of their results. The sources of the instability are different in both cases.

During portfolio optimization we must be aware that piece-wise linearity risk measures provide fast optimization, however the iso-risk surfaces are polygons where the solution exists in the extremal points. As a result, even the slightest changes in the input data can cause a jump from one extremal point to another. This can cause increased noise for piece-wise linear risk measures. In case of Expected Shortfall and Maximum Loss we must also consider that during optimization only small part of the input data is considered because our goal is to create a portfolio that minimizes extreme losses. Discarding the majority of the data leads to additional noise. This is further affected by the fact that optimization of these risk measures can only be performed with a specific probability as a function of portfolio and sample size.

Considering operational risk, the source of instability is the original Basel II regulatory definition because all institutions come face to face with the extrapolation problem regardless of the size of their operational risk database. While writing the essay it has become clear that there is no clear (and simple) answer to the question of whether institutions with small or large sample size can give a better estimate about their capital requirements, because this depends on several factors.

Overall I would like to call the attention of the reader to the fact that naive application of financial models without understanding the nature of the problem can result in delusions regarding real risk, which can be multiples of those indicated by the models.

The essay is built around **three primary research questions**:

**1. Does the extrapolation problem relating to operational risk capital requirement affect institutions with small and large samples similarly?**

- a) Accordingly, we have shown that using log-normal single loss distribution (given the expected value of frequency distribution) capital requirement is similarly log-normally distributed. This emphasizes the empirical fact that errors made during capital requirement calculation are not symmetrical, but skewed right. Besides the verification of empiricism, the importance of this derivation is the significantly faster method of distribution cognizance compared to the simulation based method (with the specific model conditions).

- b) We have shown in the introduced model environment that in case of less heavy tail (low  $\sigma$ ) single loss distribution estimation error (VaR distribution) decreases monotonously as a function of frequency, while in case of increasing  $\sigma$  the early decrease halts and starts to increase again with higher  $\lambda$  (expected frequency value) values. All this presents an interesting situation for capital estimation. Not only the amount of data the institution has is important, but also how heavy tail the loss-generating process. Their collaboration determines how precisely the institution can estimate its operational risk capital requirement.

As mentioned earlier, two forces face each other. On one hand, increasing frequency favors estimation, since greater sample means we can estimate loss distribution with lower error; on the other hand greater frequency requires further quantile extrapolation. It seems that increasing sampling size in case of a less risky process (low  $\sigma$ ) can balance the uncertainty caused by extrapolation, while this does not hold for heavy tail processes. As a result, paradoxically, in case of heavy tail processes increased sample size means greater uncertainty for capital estimation!

Naturally the situation is a bit more complicated since increasing standard deviation can still be negligible compared to the capital, so it is worth looking at relative standard deviation as well. During the examination of relative standard deviation we have experienced its monotonous decreasing characteristic. This means that banks with more loss events are in relatively better positions – regardless of distribution, because the error relative to the order of magnitude of the capital decreases monotonously. So the increasing sample size compensates for the disadvantages from the extrapolation (for the specific model conditions)!

- c) We have shown in our log-normal single loss distribution model (with certain assumptions) that VaR distribution is log-normal and its expected value and distribution both depend on the sample size ( $n$ ). Institutions with a few number of losses are expected to more drastically overestimate their operational risk capital as those with large sample. This also supports the logic that the effect of the extrapolation problem is not the same for all institutions. The less data an institution has, the more they will overestimate their operational risk requirement.
- d) We have subjected the HUNOR operational risk management database to a short analysis. The analysis has shown that the losses from specific event categories follow a heavy tail distribution. Additionally, the log-normal distribution does not properly fit at the tails and typically underestimates the

risk. We have shown that even assuming the same distribution, specific parameter estimation methods can result in capital requirements varying by orders of magnitude!

- e) Examination of the empirical data (generalized Pareto distribution fitting) has shown that the MLE estimation function is very sensitive to identical sample elements, and their management is essential.
- f) We have examined the significance of external data in a simulated environment on capital calculation, including how the generalized Pareto distribution (distribution of external data) parameters, the average number of losses of the institution, the size of the external database, and the estimation methods (MLE, PWM) affect the institution's capital requirements. In regards, we have come to the following conclusions.
  - We have experienced the (asymptotic) unbiasedness of the MLE estimation, while the PWM estimation function likely underestimates the real tail index parameter. The rate of bias is greater (for PWM) as it approaches the real parameter (tail index) 1.
  - The tail index MLE estimation is symmetrical, while the PWM is right-skewed.
  - The PWM tail index estimation is bounded from above, while the MLE is not.
  - The closer the  $\xi$  (tail index) value is to 1, the more MLE and PWM estimates differ.
  - In case of tail index MLE estimation there is a roughly 5% probability that the estimation will yield an infinite expected value distribution (in case of  $\xi=0,9$ ).
  - The PWM estimation function overestimates the scale parameter, while the MLE yields an (asymptotically) unbiased estimation.
  - The 90% confidence interval for the scale parameter is wider for PWM.
  - The difference is moderate between the scale parameter distributions for the specific estimation functions, supported by the effect of the parameter estimation error on the capital.
  - The greater  $\beta$  (scale parameter) means greater standard deviation, but relative standard deviations are similar.
  - The PWM estimation function likely underestimates the capital requirement, while the MLE estimation is asymptotically unbiased.
  - The difference between the lower and upper value of the capital requirement estimated at 90% confidence level is approximately 1 order of magnitude, even with moderate tail index!

- In case of capital requirement minimum and maximum estimated with 98% accuracy, there were thirty-fold differences! The quotient of the maximum and the minimum is always less for MLE estimation than with the PWM method. However, it is apparent that in case of  $\xi=0.9$  the upper (0.99) quantile of the PWM estimation VaR distribution is less than the same VaR quantile for the MLE estimation. The reason for this may be the phenomenon observed during the analysis of  $\xi$  – the maximum value of  $\xi$  using PWM estimation was 1, so PWM estimation never resulted a  $\xi$  value with an infinite mean GPD distribution.
- The closer the  $\xi$  value is to 1, the more varying the VaR distributions resulting from specific estimation functions.
- The relative standard deviation of the capital requirement increases as  $\xi$  approaches 1. In case of PWM it is worth noting that for  $\xi=0.9$  standard deviation increases to multiples of the expected value!
- Institution with greater frequency is not necessarily in a better position than ones with smaller frequency. By increasing annual frequency, VaR relative standard deviation (standard deviation / VaR quotient) increased for both MLE and PWM, for all  $\xi$  values! Simultaneously the rate of relative standard deviation is less for MLE estimation, for high  $\xi$  value, by almost an order of magnitude. Ranges normalized by the expected value, which includes the VaR with 98% probability, increased for both MLE and PWM estimation when frequency increased! Please note that compared to an institution with less number of data, an institution with a lot of internal data faces the clearly negative effects of the extrapolation problem. Due to the greater internal sample, in order to determine the VaR, it must extrapolate farther to the tail of the distribution than the institution with less internal data, while both have the same samples (external database) available at the tail of the distribution! This corrupts estimation quality.

We have shown in a self-created model environment that from an institution's internal sample aspect, institutions with larger sample have an advantage, because the VaR relative error decreases as sample size increases. Whereas the size of the external database is fixed (all participants have similar quantity of external data) and since the extrapolation effect is smaller for institutions with less data, its negative effects are better dampened by the external sample of the same number than for institutions with significant amounts of data where even the significant external sample is not enough for stable estimation.



Overall, even in our self-created and simplified model environment we cannot give a clear-cut answer to the research question. This means that in reality specific institutions must create their operational risk capital requirement model by carefully examining and understanding the circumstances (external-internal sampling size, loss generating process properties, etc.), and making a number of unique decisions. While the actuarial methods (LDA) provided for the problem promised fast and comfortable solutions, it has become more and more evident that there is no simple answer to the above mentioned problem.

**2. To what extent does the risk measure influence the solution of portfolio selection? In other words, how sensitive are specific risk measures to the estimation errors during optimization?**

Since standard deviation and absolute deviation results are already available in scientific literature, we primarily analyzed the Expected Shortfall and its special case – maximum loss. We have seen that the analyzed risk measures have different sensitivity levels to noise. We have observed that the Expected Shortfall and its special case, maximum loss, experience the greatest estimation error, surpassing the experiences from AD and variance. We have seen great portfolio weight fluctuation even with variance optimized portfolios, but these fluctuations were even more significant for the other observed risk measures (AD, ES, ML).

The weight instability is clear if we do not use overlapping samples. Overlapping samples are autocorrelated, so in such cases the optimum changes slowly, while in reality the portfolio get stuck in a suboptimal situation.

One of the primary messages of the analysis is that in case of alternative risk measures more information is needed (longer time series) to make them more competitive with the variance. One of the reasons for greater information demand is the piece-wise linearity of the risk measures (AD, ES, ML). Linearity provides great calculation advantages, but we have seen that it simultaneously leads to instability: sample change leads to results “jumping”, which manifests in the increased fluctuation of portfolio weights. In addition, the ES and ML risk measures only utilize only a section of the available time series – the extremes. This means that optimization from a smaller sample is possible with greater estimation error.

It is a surprise in case of ES that  $q_0$ , which measures the error during optimization, does not behave monotonously against  $\beta$  (threshold defined in probability).

**3. Is there optimal portfolio for every case? Or, is there a solution for portfolio optimization?**

In the case of Expected Shortfall and its extremal special case, maximal loss, we have encountered particularly strong sample to sample fluctuations and a surprising feasibility problem: although these are convex (in fact, coherent) risk measures, in some samples they are not bounded, and then the optimization over the weights does not have a solution. This feasibility problem may be solved by applying additional conditions, such as excluding short selling, but this means a linear programming task different from the original.

The probability of solution depends on the parameters: for maximum loss can be derived analytically, while in case of Expected Shortfall we have measured the boundary where the solution exists with a probability of 1.

## *Rövidítésjegyzék*

AD: Absolute Deviation

AD: Anderson-Darling teszt / teszt statisztika

AMA: Advanced Measurement Approach

ASA: Alternative Standardised Approach

BIA: Basic Indicator Approach

CRD: Capital Requirement Directive

ES: Expected Shortfall

ET: Event Type

FFT: Fast Fourier Transform

GoF: Goodness of fit

GPD: Generalized Pareto distribution

HUNOR: HUNgarian Operational Risk database

KS: Kolmogorov–Smirnov teszt / teszt statisztika

LDA: Loss Distribution Approach

MC: Monte Carlo szimuláció

ML: Maximal Loss

MLE: Maximum-Likelihood Estimation

O: Nagy ordó

PWM: Probability Weighted Moments

SLA: Single Loss Approximation

SMEF: Sample Mean Excess Function

MEF: Mean Excess Function

TSA: The Standardised Approach

VaR: Value at Risk

## *Hivatkozások*

200/2007. (VII. 30.) Korm. r. [2007]: A működési kockázat kezeléséről és tőkekövetelményéről. Magyar Közlöny. 101.  
<http://www.magyarokzlony.hu/nkonline/MKPDF/hiteles/MK07101.pdf>

Acerbi, C., 2004: Coherent Representations of Subjective Risk Aversion. Journal of Banking and Finance. Elsevier., in Giorgio Szegő 2004: Risk Measures for the 21st Century. Wiley, New York.

Acerbi, C., Tasche, D. [2002]: On the coherence of Expected Shortfall. Journal of Banking & Finance. 26. 1487-1503.

Ahn, S., Fessler, J. A., 2003: Standard Errors of Mean, Variance and Standard Deviation Estimators. EECS Department, The University of Michigan.

Algorithmics, 2002: Scenario, risk/reward - risk/risk optimization. Tech. rep., Algorithmics Inc.

Arató, M., 1997: Általános biztosításmatematika. ELTE, Eötvös Kiadó, Budapest.

Artzner, P., Delbaen, F., Eber, J.M., Heath, D. [1997]: Thinking coherently. Risk. 10. 68-71.

Aue, F., Kalkbrenner, M. [2006]: LDA at work: Deutsche Bank's approach to quantifying operational risk. Journal of Operational Risk. Volume 1. Number 4.

Bachelier, L., 1900: Théorie de la spéculation. Ph.D. thesis.

Bai, Z. D. [1993]: Convergence rate of expected spectral distributions of large random matrices. II. Sample covariance matrices. The Annals of Probability. 21. 649-672.

Balanda, K. P., MacGillivray, H. L. [1988]: Kurtosis: A Critical Review. The American Statistician. 42:2. 111-119.

Basel Committee on Banking Supervision, 2001: Consultative Document, Operational Risk, Supporting Document to the New Basel Capital Accord, Issued for comment by 31 May 2001. <http://www.bis.org/publ/bcbsca02.pdf>.

Basel Committee on Banking Supervision, 2003: Consultative Document, The New Basel Capital Accord, Issued for comment by 31 July 2003. <http://www.federalreserve.gov/boarddocs/press/bcreg/2003/20030430/attachment2.pdf>.

Basel Committee on Banking Supervision, 2008: Results from the 2008 Loss Data Collection Exercise for Operational Risk. <http://www.bis.org/publ/bcbs160.htm>.

Basel Committee on Banking Supervision, 2011: Operational Risk – Supervisory Guidelines for the Advanced Measurement Approaches. <http://www.bis.org/publ/bcbs196.pdf>.

Baud, N., Frachot, A., Roncalli, T., 2002: How to Avoid Over-estimating Capital Charge for Operational Risk?. Groupe de Recherche Op'erationnelle, Cr'edit Lyonnais. France.

Böcker, K., Klüppelberg, C. [2005]: Operational VAR: a closed-form approximation. Risk. 2005. 90–93.

Böcker , K., Sprittulla, J. [2006]: Operational VAR: meaningful means. Risk Magazine. 12.

Chan, L., Karceski, J., Lakonishok, J. [1999]: On portfolio optimization: Forecasting covariances and choosing the risk model. Reviews of Financial Studies 12, 937–974.

Chernobai, A., Rachev, S.T. [2006]: Applying Robust Methods to Operational Risk Modelling. Journal of Operational Risk. 1. 27-41.

Committee of European Banking Supervisors, 2006: Quantitative Impact Study 5, Overview on the Results of the EU countries. <http://www.c-eps.org/qis5.htm>.

Cope, E. W. [2011:] Penalized likelihood estimators for truncated data. *Journal of Statistical Planning and Inference*. 01/2011.

Cope., E. W., Mignola, G., Antonini G., Ugoccioni. R. [2009]: Challenges and pitfalls in measuring operational risk from loss data. *Journal of Operational Risk*. Volume 4. Number 4. 3-27.

Cruz, Marcelo G., 2002: *Modeling, Measuring, and Hedging Operational Risk*. John Wiley & Sons, Inc., New York.

Dahen, H., Dionne, G. [2008]: Scaling Models for the Severity and Frequency of External Operational Loss Data. *Journal of Banking and Finance*. Volume 34. Issue 7. 1484-1496.

Degen, M. [2010]: The calculation of minimum regulatory capital using single-loss approximations. *The Journal of Operational Risk*. Volume 5. Number 4. 3-17.

Delbaen, F. [2000]: *Coherent risk measures on general probability spaces*. Eidgenössische Technische Hochschule, Zürich.

Dickinson, J.P., [1974]: The reliability of estimation procedures in portfolio analysis. *Journal of Financial and Quantitative Analysis* 9, 447–462.

Elton, E.J., Gruber, M.J., [1973]: Estimating the dependence structure of share prices – Implications for portfolio selection. *Journal of Finance* 28, 1203–1232.

Elton, E.J., Gruber, M.J. 1995 *Modern Portfolio Theory and Investment Analysis*. J. Wiley and Sons.

Eun, C., Resnick, B. [1984]: Estimating the correlation structure of international share prices. *Journal of Finance* 39, 1311–1324.

Embrechts, P., Furrer H., Kaufmann R. [2003]: Quantifying regulatory capital for operational risk. *Derivatives Use, Trading and Regulation*. 9. 217-233.

Embrechts, P., Klüppelberg, C., and Mikosch, T., 1997: Modelling extremal events for insurance and finance. Springer, Berlin.

Ergashev, B., Pavlikov, K., Uryasev, S., Sekeris, E., 2012: Estimation of Truncated Data Samples in Operational Risk Modeling, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2193493](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2193493)

Frachot, A., Moudaoulaud, O., and Roncalli, T. [2004]: Loss distribution approach in practice. M. K. Ong (Ed.), The Basel. 369-398.

Frachot, A., Georges, P., Roncalli, T., 2001: Loss Distribution Approach for operational risk, Crédit Lyonnais, Paris, France.

Frachot, A., Roncalli, T., Salomon, E. 2004: The Correlation Problem in Operational Risk, Crédit Lyonnais, Paris, France.

Frankfurter, G., Phillips, H., Seagle, J., [1971] Portfolio selection: The effects of uncertain means, variances and covariances. Journal of Financial and Quantitative Analysis 6, 1251–1262.

Gáll, J., Nagy, G. [2007]: A működési kockázat veszteségeloszlás-alapú modellezése (Loss Distribution Approach, LDA). Hitelintézeti Szemle. 2007/4. 152-169.

Gáll, J., Pap, Gy. [2005]: Hasznosság-alapú portfóliómenedzsment. Jegyzet. Mobidiak-Debreceni Egyetem (<http://mobidiak.inf.unideb.hu>).

Gruber, M. J., Elton E. J., 1995. Modern portfolio theory and investment analysis. J. Wiley and Sons, USA.

Hajnal, B., Kállai, Z., Nagy, G. [2007]: Működési kockázati önértékelések veszteségeloszlás-alapú modellezése. Hitelintézeti Szemle. 2007/5. 506-519.

Heath, P., Artzner, P., Delbaen, F., Eber, J-M. [1999]: Coherent measures of risk. Mathematical Finance. 9. 203-228.



Jobson, J.D., Korkie, B.M. [1980]: Estimation for Markowitz efficient portfolios. *Journal of the American Statistical Association* 75, 544–554.

Kiefer, J. [1967]: On Bahadur's representation of sample quantiles, *Annals of Mathematical Statistics*. Volume 38. Number 5. 1323–1342.

Klugman, S., Panjer, H. and Willmot, G., 2004: *Loss Models, From Data to Decision*. Wiley, Hoboken, New Jersey.

Kondor, I., Nagy, G., Pafka, Sz., Karádi, R., 2006: Portfolio selection in a noisy environment using absolute deviation as a risk measure. In Takayasu, H., 2006: *Practical Fruits of Econophysics*. Springer, Tokyo.

Kondor, I., Szepessy, A., Újvárosi, T., 2004: Concave risk measures in international capital regulation. In: Szegő, G., 2004: *Risk measures for the 21st century*. John Wiley & Sons.

Kondor, I., Varga-Haszonits, I. [2010]: Instability of portfolio optimization under coherent risk measures. *Advances in Complex Systems*. 13. 425-437.

Konno, H., Yamazaki, H. [1991]: Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. *Management Science*. 37. 519-531.

Laloux, L., Cizeau, P., Bouchaud, J-P., Potters, M. [1999]: Noise dressing of financial correlation matrices. *Physical Review Letters*. 83. 1467-1470.

Laloux, L., Cizeau, P., Potters, M., Bouchaud, J-P. [2000]: Random matrix theory and financial correlations. *International Journal of Theoretical and Applied Finance*. 3. 391-397.

Lynch, S. M. [2007]: *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. Springer-Verlag, New York.

Marchenko, V., Pastur, L. [1967]: Distribution of eigenvalues for some sets of random matrices. *Math. USSR-Sbornik* 1. 457-486.

Markowitz, H. [1952]: Portfolio selection. *The Journal of Finance*. Vol. VII. No. 1. 77-91.

Markowitz, H., [1959]. *Portfolio selection: efficient diversification of investments*. J. Wiley and Sons, New York.

Mattheiss, T.H., Schmidt, B. K. [1977]: The probability that a random polytope is bounded. *Mathematics of Operations Research*. 2. 292-296.

Merton, R. C. [1972]: An analytic derivation of the efficient portfolio frontier. *Journal of Financial and Quantitative Analysis*. 7. 1851-1872.

Mignola, G., Ugocioni, R. [2006]: Sources of uncertainty in modelling operational risk losses. *Journal of Operational Risk*. Volume 1(2). 33–50.

Moscadelli, M., 2004: The modeling of operational risk: experience with the analysis of the data collected by the Basel committee. Technical Report 517. Banca d'Italia.

Nagy, G., Povilaitis, K. [2009]: A működési kockázatok veszteségeloszlás alapú modellezésének lehetőségei. *Hitelintézeti Szemle*. 2009/2. 152-169.

Noh, J. D. [2000]: Model for correlations in stock markets. *Physical Review*. 61. 5981-5982.

Nowak, M.A., Jurkiewicz, J., Burda, Z. [2003]: Is Econophysics a solid science? *Acta Physica Polonica*. 34. 87-132.

Ong, M. K., 2006: *The Basel Handbook: A Guide for Financial Practitioners*, Ong, M. (ed). Risk Books. London.

Pafka, Sz., Kondor, I. [2002]: Noisy covariance matrices and portfolio optimisation. *European Physical Journal*. B 27. 277-280.

Pafka, Sz., Kondor, I. [2003]: Noisy covariance matrices and portfolio optimisation II. *Physica*. 319. 487-494.

Pafka Sz., Potters M., Kondor I. [2004], Exponential Weighting and Random-Matrix-Theory Based Filtering of Financial Covariance Matrices for Portfolio Optimization, <http://arxiv.org/pdf/cond-mat/0402573>

Panjer, H. [1981]: Recursive evaluation of compound distributions. *Astin Bulletin*. Volume 12. 22-26.

Panjer, H., 2006: *Operational Risk, Modeling Analytics*. Wiley, Hoboken, New Jersey.

Panjer, H., Lutek, B. [1983]: Practical Aspects of Stop-Loss Calculations. *Insurance: Mathematics and Economics*. Volume 2. 159-177

Panjer, H., Wang, S. 1993: On the stability of recursive formulas. *ASTIN Bulletin*, University of Waterloo, Ontario, Canada.

Panjer, H. and Willmot, G. [1986]: Computational Aspect of Recursive Evaluation of Compound Distributions. *Insurance: Mathematics and Economics*. Volume 5. 113-116.

Panjer, H. and Willmot, G., 1992: *Insurance Risk Models*. Society of Actuaries, Chicago.

Papp, G., Kondor, I., Pafka, Sz., Nowak, M.A. [2005]: Random matrix altering in portfolio optimization. *Acta Physica Polonica*. B 36. 2757-2765.

Patchin, C., Carey M., 2012 : *Risk Assessment in Practice*. Committee of Sponsoring Organizations of the Treadway Commission

Pénzügyi Szervezetek Állami Felügyelete, 2006: *Validációs kézikönyv*

Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral L.A.N., Stanley, H.E. [1999]: Universal and nonuniversal properties of cross correlations in financial time series. *Physical Review Letters*. 83. 1471-1474.

Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral L.A.N., Guhr, T., Stanley, H.E. [2002]: Random matrix approach to cross correlations in financial data. *Physical Review*. 65. 66-126.

Rockafellar, R. T., Uryasev, S. [2000]: Optimization of Conditional Value-at-Risk. *The Journal of Risk*. 2. 21-41.

Rockafellar, R. T., Uryasev, S. [2002]: Conditional Value-at-Risk for general loss distributions. *Journal of Banking & Finance*. 26. 1443-1471.

Roehr, A. [2002]: Modelling Operational Losses. *Algo Research Quarterly* 53. 5/2. 53-64.

Ruppert, D. [1987]: What is Kurtosis? An Influence Function Approach. *The American Statistician*. 41/1. 292-307.

Rychlik, I., Rydén, J., 2006: *Probability and risk analysis: an introduction for engineers*. Springer, Berlin.

Samad-Khan, A. [2005]: Why COSO is Flawed. *Operational Risk*. 1. 1-6.

Samad-Khan, A. [2008]: Modern Operational Risk Management. *Emphasis*. 2008/2. 26-29.

Schmidt, B.K., Mattheiss, T. [1977]: The probability that a random polytope is bounded. *Mathematics of Operations Research* 2, 292–296.

Shin, H. S., Keating, C., Goodhart, C., Danielsson, J. [2001]: An academic response to Basel II. *FMG – ESRC*. 130.

Simaan, Y. [1997]: Estimation risk in portfolio selection: The mean variance model versus the mean absolute deviation model. *Management Science*. 43. 1437-1446.

Szabolcsné Nikházy, O. [2009]: A magyar bankszektor működési kockázatai a pénzügyi válság tükrében. *Hitelintézeteti Szemle*. 2009/5.

Szegő, G., 2004: *Risk measures for the 21st century*. Wiley, Somerset, New Jersey, USA.

Todd, M. [1991]: Probabilistic models for linear programming. *Mathematics of Operations Research*. 16. 671-693.

Wigner, E. P. [1958]: On the distribution of the roots of certain symmetric matrices. *The Annals of Mathematics*. 67. 325-327.

Young, M. R. [1998]: A minimax portfolio selection rule with linear programming solution. *Management Science*. 44. 673-683.