

LOLBERT TAMÁS

**STATISZTIKAI ELJÁRÁSOK ALKALMAZÁSA
AZ ELLENŐRZÉSBEN, KÜLÖNÖS TEKINTETTEL
A PÉNZÜGYI ELLENŐRZÉSRE**

Budapesti Corvinus Egyetem, Matematika Tanszék
Állami Számvevőszék

Témavezető: dr. Hunyadi László egyetemi tanár

BUDAPESTI CORVINUS EGYETEM
KÖZGAZDASÁGTANI DOKTORI ISKOLA

**Statisztikai eljárások alkalmazása
az ellenőrzésben, különös tekintettel
a pénzügyi ellenőrzésre**

Ph.D. értekezés

Lolbert Tamás

Budapest, 2008.

Tartalomjegyzék

TARTALOMJEGYZÉK	i
ÁBRAJEGYZÉK	iii
TÁBLÁZATOK JEGYZÉKE	iv
KÖSZÖNETNYILVÁNÍTÁS	v
BEVEZETÉS	1
I. RÉSZ: AZ ELLENŐRZÉS	7
AZ ELLENŐRZÉS TÖRTÉNELMI GYÖKEREI, FŐBB CÉLJAI	7
AZ ELLENŐRZÉS FOLYAMATA	17
A MINTAVÉTEL SZEREPE	19
STATISZTIKAI ÉS NEM STATISZTIKAI MINTAVÉTELI ELJÁRÁSOK	20
BIZONYOSSÁG, PONTOSSÁG, LÉNYEGESSÉG	21
A NULLHIPOTÉZIS SZEREPE	26
AZ ELLENŐRZÉSI KOCKÁZAT JELENTÉSE, A „MATEMATIKAI MODELL” KRITIKÁJA	27
II. RÉSZ: ELEMZŐ ELJÁRÁSOK	33
AZ ELEMZŐ ELJÁRÁSOK CÉLJA ÉS FŐBB JELLEGZETESSÉGEI	33
AZ ELEMZŐ ELJÁRÁSOK ALKALMAZÁSA ALAPVETŐ ELJÁRÁSKÉNT	34
AZ ELEMZŐ ELJÁRÁSOK ALKALMAZÁSA A KOCKÁZATELEMZÉSBEN	36
A KOCKÁZAT DIMENZIÓI, MÉRÉSE	36
AZ ELLENŐRZÉSI SZAKMÁBAN HAGYOMÁNYOSNAK TEKINTHETŐ KOCKÁZATELEMZÉSI ELJÁRÁS	39
AZ ELLENŐRZÉSBEN HASZNÁLT KOCKÁZATELEMZÉSI MÓDSZEREK TOVÁBBFEJLESZTÉSI LEHETŐSÉGEI	42
<i>Az eredendő és a kontrollkockázat kapcsolata</i>	<i>43</i>
<i>A kockázati tényezők aggregálási módja</i>	<i>44</i>
A SZÁMJEGYELEMZÉS, MINT SPECIÁLIS KOCKÁZATELEMZÉSI MÓDSZER	44
A BENFORD-TÖRVÉNY TÖRTÉNETE	46
LEHETSÉGES MAGYARÁZATOK A TÖRVÉNYSZERŰSÉGRE	46
A BENFORD-TÖRVÉNY EGY ÚJ MAGYARÁZATA	50
ALKALMAZÁSOK	56
A BENFORD-TÖRVÉNY ÉRVÉNYESÜLÉSÉNEK STATISZTIKAI TESZTELÉSE	60
KÖVETKEZTETÉSEK, A SZÁMJEGYELEMZÉS ALKALMAZÁSI KORLÁTAI	66

III. RÉSZ: ARÁNYBECSLÉSRE VISSZAVEZETHETŐ ELLENŐRZÉSI FELADATOK69

A KONTROLLTESZTELÉS IDŐSZERŰSÉGE: SOX 404.....	70
A KONTROLLTESZTELÉS ALKALMAZÁSA AZ ÁLLAMI SZÁMVEVŐSZÉK MUNKÁJÁBAN	71
PONTBECSLÉSEK.....	73
INTERVALLUMBECSLÉS, HIPOTÉZISVIZSGÁLAT (EGYLÉPCSŐS MINTAVÉTEL)	76
<i>A hibamentes mintából levonható következtetések</i>	76
<i>A mintavételi statisztikához kapcsolódó becslések</i>	78
<i>Bayesi szemléletű becslések</i>	82
<i>Egy vegyes becslési módszer</i>	84
AZ EGYLÉPCSŐS BECSLÉSI ELJÁRÁSOK ÉRTÉKELÉSE.....	85
<i>Az értékelés módszere</i>	85
<i>Az első szakasz eredményei</i>	87
<i>A második szakasz eredményei</i>	88
<i>Következtetések</i>	89
INTERVALLUMBECSLÉS, HIPOTÉZISVIZSGÁLAT (SZEKVENCIAÁLIS MINTAVÉTEL)	90
<i>Az önkormányzatok átfogó ellenőrzéséhez kidolgozott mintavételi módszer</i>	93
<i>A módszer alkalmazási tapasztalatai a 2007. év első félévében</i>	97

IV. RÉSZ: ÉRTÉKÖSSZEG-BECSLÉSRE VISSZAVEZETHETŐ ELLENŐRZÉSI FELADATOK....101

A KÖNYVVIZSGÁLATBAN ELŐFORDULÓ POPULÁCIÓK FŐBB STATISZTIKAI JELLEMZŐI.....	103
A VALÓS ÉRTÉK PONTBECSLÉSE	104
A VALÓS ÉRTÉK INTERVALLUMBECSLÉSE.....	105
<i>Egy EV mintán alapuló becslés</i>	106
<i>MUS mintán alapuló becslések</i>	107
KÖVETKEZTETÉSEK	121
A MULTINOMIÁLIS BECSLÉS ÉS A STRINGER-FÉLE KÉPLET ELEMZÉSE.....	122

ÖSSZEGZÉS.....131

IRODALOMJEGYZÉK135

SAJÁT PUBLIKÁCIÓK JEGYZÉKE141

FÜGGELÉK.....143

Ábrajegyzék

1. ÁBRA: A LOPÁSI KOCKÁZAT ELLEN MŰKÖDTETETT KONTROLLK OPTIMÁLIS SZINTJE	10
2. ÁBRA: A LÉNYEGES HIBA BEKÖVETKEZÉSE	28
3. ÁBRA: AZ ELLENŐRZÉS KOCKÁZATI TÉNYEZŐINEK KAPCSOLATA A VÁLASZTOTT ELLENŐRZÉSI MÓDSZERREL ..	31
4. ÁBRA: AZ ADATOK VIZUÁLIS MEGJELENÍTÉSE AZ IDEA SZOFTVERBEN.....	61
5. ÁBRA: A 4. ÁBRA ALAPADATAI AZ IDEA SZOFTVERBEN	64
6. ÁBRA: A SZEKVENCIALIS TESZT LEHETSÉGES FORGATÓKÖNYVEI.....	92
7. ÁBRA: A MÓDOSÍTOTT SZEKVENCIALIS TESZT LEHETSÉGES FORGATÓKÖNYVEI	95
8. ÁBRA: AZ EGYLÉPCSŐS ÉS A SZEKVENCIALIS TESZT ÖSSZEHASONLÍTÁSA	96
9. ÁBRA: TÉTELEK A SZISZTEMATIKUSAN KIVÁLASZTOTT MUS MINTÁBAN	108
10. ÁBRA: CELLÁS MÓDSZERREL KIVÁLASZTOTT TÉTELEK A MUS MINTÁBAN	109
11. ÁBRA: A SZENNYEZETTSÉG MÉRÉSE.....	110
12. ÁBRA: A KÖNYV SZERINTI ÉRTÉK SZENNYEZETTSÉGE	113
13. ÁBRA: A STRINGER-FÉLE FELŐ HATÁR ASZIMPTOTIKUS MEGBÍZHATÓSÁGA	117
14. ÁBRA: AZ EXTRÉMEBB KIMENETEK STEP DOWN S HALMAZÁNAK EGY METSZETE.....	124
15. ÁBRA: A STEP DOWN S HALMAZOKKAL DEFINIÁLT MULTINOMIALIS BECSLÉS ELUTASÍTÁSI TARTOMÁNYA ...	126
16. ÁBRA: AZ M3 BECSLÉS MEGJELENÍTÉSE EXCEL BEN	144
17. ÁBRA: A B1 BECSLÉS PARCIÁLIS ÉS TELJES MEGBÍZHATÓSÁGA	145
18. ÁBRA: AZ ELEMZÉS I. OUTPUTJA A (200/30/5%) FORGATÓKÖNYV ESETÉN.....	146
19. ÁBRA: AZ ELEMZÉS II. OUTPUTJA A (200/30/5%) FORGATÓKÖNYV ESETÉN	147
20. ÁBRA: AZ ELEMZÉS I. OUTPUTJA A (200/50/5%) FORGATÓKÖNYV ESETÉN.....	148
21. ÁBRA: AZ ELEMZÉS II. OUTPUTJA A (200/50/5%) FORGATÓKÖNYV ESETÉN	149
22. ÁBRA: AZ ELEMZÉS I. OUTPUTJA A (200/150/5%) FORGATÓKÖNYV ESETÉN.....	150
23. ÁBRA: AZ ELEMZÉS II. OUTPUTJA A (200/150/5%) FORGATÓKÖNYV ESETÉN	151
24. ÁBRA: AZ ELEMZÉS I. OUTPUTJA A (600/30/5%) FORGATÓKÖNYV ESETÉN.....	152
25. ÁBRA: AZ ELEMZÉS II. OUTPUTJA A (600/30/5%) FORGATÓKÖNYV ESETÉN	153
26. ÁBRA: AZ ELEMZÉS I. OUTPUTJA A (600/50/5%) FORGATÓKÖNYV ESETÉN.....	154
27. ÁBRA: AZ ELEMZÉS II. OUTPUTJA A (600/50/5%) FORGATÓKÖNYV ESETÉN	155
28. ÁBRA: AZ ELEMZÉS I. OUTPUTJA A (600/150/5%) FORGATÓKÖNYV ESETÉN.....	156
29. ÁBRA: AZ ELEMZÉS II. OUTPUTJA A (600/150/5%) FORGATÓKÖNYV ESETÉN	157
30. ÁBRA: AZ ELEMZÉS I. OUTPUTJA AZ (1500/50/5%) FORGATÓKÖNYV ESETÉN.....	158
31. ÁBRA: AZ ELEMZÉS II. OUTPUTJA AZ (1500/50/5%) FORGATÓKÖNYV ESETÉN	159
32. ÁBRA: AZ ELEMZÉS I. OUTPUTJA AZ (1500/200/5%) FORGATÓKÖNYV ESETÉN.....	160
33. ÁBRA: AZ ELEMZÉS II. OUTPUTJA AZ (1500/200/5%) FORGATÓKÖNYV ESETÉN	161

Táblázatok jegyzéke

1. TÁBLÁZAT: AZ ELLENŐRZÉS SORÁN POTENCIÁLISAN ELŐFORDULÓ NÉGY (KÉT) SZITUÁCIÓ	26
2. TÁBLÁZAT: A HIPOTÉZISVIZSGÁLAT SORÁN POTENCIÁLISAN ELŐFORDULÓ NÉGY SZITUÁCIÓ	26
3. TÁBLÁZAT: AZ ELSŐ SZÁMJEGYEK RELATÍV GYAKORISÁGA TÍZES SZÁMRENDSZERBEN	45
4. TÁBLÁZAT: AZ „1” SZÁMJEGY MEGJELENÉSE KEZDŐ SZÁMJEGYKÉNT A SZORZÓTÁBLÁBAN	47
5. TÁBLÁZAT: A MÉTERBEN MÉRT EMBERI TESTMAGASSÁG KÜLÖNBÖZŐ HATVÁNYAINAK ELSŐ-SZÁMJEGY ELOSZLÁSA	56
6. TÁBLÁZAT: A MAD TESZT NIGRINI ÁLTAL JAVASOLT KRITIKUS ÉRTÉKEI	63
7. TÁBLÁZAT: MIKOR HASZNÁLHATÓ ÉS NEM HASZNÁLHATÓ A SZÁMJEGYELEMZÉS	67
8. TÁBLÁZAT: A NULLA MINTABELI HIBA VALÓSZÍNŰSÉGE A SOKASÁGI HIBA FÜGGVÉNYÉBEN	76
9. TÁBLÁZAT: A TESZTELÉSHEZ HASZNÁLT FORGATÓKÖNYVEK	86
10. TÁBLÁZAT: A MÓDOSÍTOTT SZEKVENCIAÁLIS MÓDSZER ALKALMAZÁSÁNAK TAPASZTALATAI	99
11. TÁBLÁZAT: AZ „A” TÍPUSÚ MÉRLEG FELÉPÍTÉSE	101
12. TÁBLÁZAT: EGY HIPOTETIKUS ÚTIKÖLTSÉG-ELSZÁMOLÁS	110
13. TÁBLÁZAT: AZ ÚTIKÖLTSÉG-ELSZÁMOLÁS TÉTELEINEK KÖNYV SZERINTI- ÉS KUMULÁLT KÖNYV SZERINTI ÉRTÉKE	111

Köszönetnyilvánítás

Amikor a Ph.D. programra jelentkeztem, azt terveztem, hogy az elektromos áram piacán kereskedett termékek árazását kutatom majd, de a sors úgy akarta, hogy kutatási témát váltsak. Akkori munkahelyemen, a Kormányzati Ellenőrzési Hivatalnál egy módszertani anyag készítésekor belefutottam az értekezés III. részében leírt aránybecslési problémába, és ezzel kapcsolatos kérdéseimmel, ötleteimmel korábbi egyetemi oktatómhoz, dr. Hunyadi Lászlóhoz fordultam. Az ő biztatására írtam le a beszélgetésen elhangzott gondolatokat, amiből megszületett az első komolyabb publikációm.

Amikor ez a cikk megjelent, már az Állami Számvevőszéknél dolgoztam, ahol az alkalmazások újabb és újabb, statisztikai szempontból „egzotikus” témákat vetettek fel: mindez oda vezetett, hogy kutatási témát váltottam. A témavezetést dr. Hunyadi László vállalta el, sokszor erején felül is foglalkozott az általam készített munkaanyagokkal, és azok publikálásra való előkészítésével. Az ő javaslatai, kérdései által sokkal közérthetőbbé vált ez az értekezés is, ami különösen nagy szó egy ennyire multidiszciplináris kutatási terület esetén.

Ezúton szeretnék tehát köszönetet mondani dr. Hunyadi Lászlónak, akinek segítsége, bátorítása nélkül ez az értekezés nem született volna meg.

Köszönet illeti még az Állami Számvevőszék felsővezetőit, akik lehetővé tették ennek az értekezésnek az elkészítését és a mögöttes módszertani kutatások lefolytatását, továbbá az Állami Számvevőszék könyvtárát is, ahol a ritka alapműveket számomra beszerezték.

Külön köszönet illeti dr. Sepsey Tamást, aki a jelen értekezést a tőle megszokott alaposággal átnézte, és több hasznos ötlettel, észrevétellel javította annak színvonalát.

Sokat tanultam azokból a beszélgetésekből, vitákból, amelyeket az ellenőrzési eljárásokkal, főként a mintavétellel, a mintákból levonható következtetésekkel kapcsolatosan folytattam kollégáimmal, különösen Bodonyi Miklóssal és Csecserits Imrénével, illetve mindazokkal, akik részt vettek valamelyik általam tartott belső továbbképzésen. Szerencsésnek érzem magam amiatt is, hogy részt vehettem a 2007–2010 közötti időszak átfogó ellenőrzési módszertanát kidolgozó munkacsoportban, és ott együtt dolgozhattam olyan szakemberekkel, akiknek felkészültsége, tapasztalata valóban példaértékű.

Nem lehet megfedkezni a Kormányzati Ellenőrzési Hivatalnál eltöltött csaknem három évről sem, ahol a gyakorlati tapasztalatok, illetve az EU-ban bevált ellenőrzési módszerek átvételével kapcsolatos szemináriumok megalapozták a későbbi elméleti tudást is.

Hálás vagyok dr. Tallos Péternek, a Corvinus Egyetem Matematika Tanszék vezetőjének, aki biztosította számomra a Tanszékét, mint hivatalos kutatóhelyet, és ezt a tanszéken kívüli témavezetőre való váltást követően is fenntartotta. Ezen kívül köszönöm a megtisztelő lehetőséget, hogy a Tanszék munkájában öt féléven keresztül oktatóként is részt vehettem.

Sokat segítettek az eredmények elérésében, és ezért köszönet illeti a Matematika Tanszékén dolgozó kollégákat – különösképpen dr. Kánnai Zoltánt – is, akiknek javaslatai, ötletei többször továbblényítették az elakadás szélén álló kutatást.

Végül pedig szeretnék köszönetet mondani volt témavezetőmnek, dr. Medvegyev Péternek, aki annak idején elvállalta eredeti kutatási témám vezetését, és pár jó ötlettel ezt a kutatást is segítette.

Budapest, 2007. április

Jelen disszertáció 2007 szeptemberében lefolytatott műhelyvitáján opponenseim (dr. Arató Miklós és dr. Kovács Árpád) és a hozzászólók közül dr. Nyikos László sok olyan hasznos észrevételt tett, amelyeket figyelembe véve ez a végleges anyag elkészülhetett. Javaslataik beépítése – megítélésem szerint – sokat javított a disszertáció színvonalán, munkájukat e sorokkal is szeretném még egyszer megköszönni.

Budapest, 2007. november

Lolbert Tamás

Bevezetés

Az emberi társadalom kialakulása egyet jelentett a szervezetek létrehozatalával. Kezdetben ez ugyan nem részvénytársaságokat, vagy minisztériumokat jelentett, de a közösségi akarat kialakítása és érvényesítése mindig is együtt járt bizonyos feladat- és hatásköröknek a társadalom egyes tagjainál való koncentrálódásával.

A szervezetek működése természetesen kockázatokat hordoz magában: olyan események bekövetkezési lehetőségét, amelyek megvalósulása hátráltathatja, korlátozhatja, vagy megakadályozhatja azon célok elérését, melyek elérésére a szervezetet létrehozták.

Szinte már a kezdetek kezdetén felismerték, hogy a szervezetek eredendő kockázatai csökkenthetők. Az első szervezeteknél ennek módja tipikusan az erőszak, illetve az erőszak kilátásba helyezése volt, mint ahogy ezt most is megfigyelhetjük a maffia jellegű szervezeteknél. Más szervezetek alapvetően a megfelelően kiválogatott, öntudatos személyi állományuknak köszönhetően tudtak úrrá lenni a kockázatokon, mint például a legtöbb egyházi, vagy egyéb jótékonyági szervezet.

Tömeges méretekben az újkor hajnalán kezdték a kockázatokat csökkenteni azáltal, hogy a szervezetet szabályok szerinti működésre készítették. Ez együtt járt a munkafolyamatok pontos leírásával, automatizmusok létrehozásával, az emberi munka gépiessé válásával, és később gépekkel való kiváltásával.

Amennyiben létre is hoztunk egy „ideális” szabályrendszert, ez nem azt jelenti, hogy az érintett szervezeti folyamatok eredendő kockázatai teljesen kiküszöbölhetőek volnának. Ilyen ideális esetben például a tízparancsolat önmagában biztosítaná, hogy nem volna bűnözés. A szabályrendszer megalkotása azonban nem feltétlenül jár annak megvalósulásával: a példánál maradva, noha több jogszabály is tiltja, mégis naponta ölnek meg, vagy rabolnak ki embereket. Amennyiben egy szabályt nem tartanak be, akkor nem érvényesül az adott szabály kockázatsökkentő hatása: a nem szabályszerű működés miatt továbbra is érvényesülhetnek mindazon kockázatok, melyeket a szabály betartásával ki lehetett volna szűrni.

Bizonyos szervezeti nagyság felett a szervezeti működés feladat-delegálás láncolatával írható le legjobban, a feladat-delegálás azonban együtt jár – az elméleti közgazdaságtanban elég alaposan kielemezett – megbízó-ügynök probléma kiteljesedésével is. Pont emiatt minden vezető elemi érdeke, hogy a lehető legtöbb információnak jusson a birtokába, például úgy,

hogy megköveteli a folyamatok dokumentálását. A szervezeti működésnek ezt a jellegzetességét nevezi a köznapí szóhasználat bürokráciának. Hiába működhetne jól egy szervezet ezek hiányában is, egy vezető sem engedheti meg magának a bizonyosság hiányát.

Az a tény, hogy a szervezeti működés minőségével, annak eredményeinek megfelelőségével kapcsolatosan gyakorlatilag senki nem rendelkezik tökéletes információval, egy új szakmát hozott létre, az auditálást. A szó eredetével kapcsolatban létezik egy, az auditorok körében közkeletű történet, melynek legismertebb változata a következőképpen hangzik:

Régen, a Római Birodalom idejében adófizetéskor meg kellett jelenni egy testület előtt, és ott be kellett számolni minden, az adófizetés szempontjából lényeges körülményről. A beszámolót hivatásos hallgatóság, úgynevezett „auditor”-ok¹ előtt kellett megtartani, akik szakértelmüknek köszönhetően nyilatkozni tudtak a beszámoló hihetőségéről.

Akár igaz, akár kitalált ez a történet, az **auditálás célja bizonyosság nyújtása**² egy állítás (pontosabban egy komplex állítás-halmaz) valódiságáról.

Fontos megjegyezni, hogy az audit tevékenység kicsit más értelmet nyer a közgazdasági modellek többségében. Azon modellek jelentős részénél, amelyekben nincs teljes információ, például információs aszimmetria miatt, általában létezik egy olyan almodell, aminek a nevében megjelenik a „with auditing” kitétel. Ezekben a modellekben az auditálás általában arról szól, hogy a megrendelő bizonyos díj ellenében biztos információhoz jut, és kevesebb az olyan (bayesi) modell, ahol a díjfizetés csak részben, valamilyen posterior valószínűségen keresztül csökkenti a bizonytalanságot, ahogyan azt egy valós audit is teszi.

Az „igazolandó” állítás típusa alapján különféle auditokról beszélhetünk, de ezek legtöbbje valamilyen módon a szervezet gazdálkodásával van összefüggésben. Ennek következtében a hazai köznyelv – de sokszor sajnos a szakmai nyelv is – egyszerűen könyvvizsgálatnak hívja az audit tevékenységet³, noha az a lehetséges auditoknak csak egy (de vitathatatlanul legnagyobb) csoportját jelenti. A könyvvizsgálat mellett léteznek más, gyakorlat szempontból jelentős audit területek is, ilyen például a minőségbiztosítási audit.

¹ Maga az auditor szó is a latin audire (hallani, hallgatni) igéből származtatható.

² Az értekezés hátralévő részében a bizonyosság szót – összhangban a hazai ellenőrzési terminológiával – nem az angol „certainty”, hanem az „assurance” értelmének megfelelően fogom használni.

³ Az audit szó angol nyelvű értelmező szótárakban rögzített jelentése „(official) examination”, ami magyarra fordítva hivatalos vizsgálat, esetleg hivatalos ellenőrzés. Tekintettel arra, hogy a szó jelenlegi értelmében angolszász közvetítéssel került a magyar nyelvbe, magyar megfelelőjeként általában az „ellenőrzés” vagy a „vizsgálat” szavakat fogom használni, a könyvvizsgálat szót pedig fenntartom az „audit of financial statements” kifejezés fordítására.

A közszféra „állításainak” igazolása, a közszféra ellenőrzése állandó szakmai kapcsolatban van a magánszféra könyvvizsgálatával, mégis önálló „alfajként” említendő. A leginkább markáns csoportot ezen belül az úgynevezett számvevőszéki ellenőrzés jelenti, amelyet az egyes államok legfőbb ellenőrző intézményei⁴, magyarul a számvevőszékek végeznek. A számvevőszékek jogköre, feladatai közel sem egységesek⁵, azonban ellenőrzési módszereik közelítése megalakulása óta programja az INTOSAI-nak, a számvevőszékek nemzetközi szervezetének.

Az auditori állítások természetesen nem lehetnek kategorikusak, nem „abszolút igazságok”, **csupán bizonyosságot nyújtanak**. A bizonyossági szint komplementere feleltethető meg a statisztikában elsőfajú hiba névvel jelzett fogalomnak, tehát azt mutatja, hogy mekkora a nullhipotézis téves elutasításának kockázata. Mivel a másodfajú hiba elkövetésének kockázatát csak közvetett módon lehet kezelni, ezért általában a nem-megfelelőséget szokás nullhipotézisként megfogalmazni, azaz a jól kontrollálható elsőfajú hiba az auditor kockázata, a másodfajú hiba pedig a vizsgált szervezeté.

Az audit lefolytatásakor az auditornak választania kell a szóba jöhető ellenőrzési eljárásokból, és ezeknek megfelelően lefolytatnia az ellenőrzést. A könyvvizsgálatban jelenleg elterjedt eljárások nem képesek megalapozottan számszerűsíteni a kockázatok nagy részét. Noha az általánosan használt elemző eljárások és a szubjektív szempontok alapján kiválasztott tételek vizsgálata rengeteg hasznos információval szolgálhat, a statisztikai mintavételi módszerek alkalmazásával sokkal jobban mérhetővé válik a téves auditori vélemény kiadásának kockázata. Nem véletlen tehát, hogy az utóbbi években egyre inkább teret nyertek a valószínűségi mintán alapuló eljárások.

A valószínűségi mintán alapuló eljárások egyik legfontosabb jó tulajdonsága, hogy számszerűsíthetővé válik az átváltás („trade-off”) a könyvvizsgálati munkaráfordítás, a kijelentések megbízhatósága, illetve a kijelentések pontossága között, legalábbis akkor, ha a munkaráfordítást a mintamérettel (az ellenőrzendő tételek számával), a megbízhatóságot pedig a konfidenciaszinttel mérjük, és eltekintünk a nem mintavételi hibák lehetőségétől. Az átváltások számszerűsítése teszi lehetővé az optimális munkaráfordítás meghatározását.

⁴ SAI, Supreme Audit Institution

⁵ Az EU régi 15 tagállamában működő számvevőszékek jogi státuszáról, feladatairól hasznos összefoglaló elemzés található egyebek mellett Nyikos et al. [2002] I. kötetében.

A világ egyre inkább abba az irányba halad, hogy ellenőrizhető, reprodukálható és módszertanilag is korrekt tevékenységet várunk el minden szakmától. Mint ahogyan az auditorok is pozitívabban ítélik meg azokat a szervezeteket, ahol a belső folyamatok egy jól átgondolt szabályozást követnek, és nem esetlegesek, éppen így az auditori tevékenység objektivitásának növekedése is előrelépésnek tekinthető.

Pont emiatt jelent súlyos problémát Magyarországon az, hogy egyes ellenőrzési alapfogalmak értelmezése nem tisztázott, még jogszabályi szinten sem.⁶ Tekintsük akár az Alkotmányt, akár egyéb törvényeket: olyan, rosszul definiált, ellenőrzésszakmai szempontból nehezen megfogható kifejezésekkel találkozunk ezekben, mint „pénzügyi-gazdasági ellenőrzés” vagy „államháztartás gazdálkodásának ellenőrzése”. Mindez azért okoz problémát, mivel a közpénzek ellenőrzésére a „megbízást” általában a jogszabályok adják (nem pedig egyedi megbízások), és – szemben egy hagyományos megbízással – ilyenkor nincs lehetősége a megbízottnak „visszakérdezni”, tisztázni, hogy pontosan mi is volna a feladat. Amellett, hogy önmagában is aggályosnak tekinthető, ha egy megbízottnak kell „megfejtienie”, mi is a feladata, közpénzből finanszírozott ellenőrzések esetén ez különösen veszélyes, hiszen az az ellenőrzésre rendelkezésre álló források felesleges, nem számonkérhető felhasználást okozhatja. Mivel ez a probléma messzire vezet, továbbá szakmai előéletem sem hatalmaz fel a probléma „megoldására”, ezért ezzel a kérdéskörrel – nevezetesen az alapfogalmak és az ellenőrzési feladatok, célok *de lege ferenda* szemléletű meghatározásával – ebben az értekezésben nem kívánok foglalkozni. A dolgozatban abból a hipotézisből fogok kiindulni, *mintha* az ellenőrzési célok, feladatok mindenkor jól definiáltak lennének.

A kutatásom tárgya ennek megfelelően leszűkül egy olyan ellenőrzési részterületre, aminek fejlesztése álláspontom szerint leginkább növelheti az ellenőrzések hatékonyságát, akár hazai, akár világviszonylatban nézzük. Ez a részterület pedig nem más, mint az angolul „audit sampling”-nek nevezett tudományterület. Az értekezés célja feltérképezni, egységes szerkezetbe foglalni, és végül a lehetőségekhez mérten statisztikailag-matematikailag is alátámasztani mindazon elemző, mintavételi és minta-kiértékelési eljárásokat, amelyek az egyes vizsgálati célokkal összefüggésben szóba jöhetnek, és korunk ellenőrzési gyakorlatában elterjedtek. Ennek érdekében először „lefordítom” a statisztika és a valószínűségszámítás nyelvére a gyakorlatban előforduló lényegesebb eljárásokat, majd ezt követően a

⁶ Erről a problémáról, nevezetesen a jogszabályi szinten is meglévő terminológiai zavarról részletes ismertetés található Nyikos et al. [2002] több fejezetében is. Ez a forrás azért tekinthető a témában különösen mérvadónak, mivel szerzője az ÁSZ egykori alelnöke, lektora az ÁSZ jelenlegi (2007) elnöke.

rendelkezésre álló matematikai-statisztikai eszköztár segítségével megvizsgálom, mennyire megalapozottak. E fő cél mellett a dolgozatnak célja még a túlmisztifikált, vagy „szokásjogon” alapuló gyakorlati eljárások „objektív” értékelése, továbbá – lehetőség szerint – új ellenőrzési módszerek kidolgozása, más tudományterületekről származó módszerek adaptálása az ellenőrzésre, valamint a meglévő módszerek továbbfejlesztése.

Az értekezés címéből is kitűnik, hogy – noha alapvetően a tágabban értelmezett pénzügyi ellenőrzés szempontjából közelíti meg a témát – megpróbálom a témakört a lehető legáltalánosabban tárgyalni, és így rámutatni, hogy a szóban forgó jelenség például nem csak egy könyvvizsgálat, vagy egy számvevőszéki ellenőrzés során, hanem minden „bizonyosságot nyújtó szolgáltatás”⁷, ad absurdum például egy élelmiszerhatósági ellenőrzés során is hasonlóképpen érvényes. Az általánosítást segíti, hogy az IFAC könyvvizsgálati standardja is tartalmaz egy ilyen, általános jellegű kitekintést minden, bizonyosságot nyújtó szolgáltatásra (International Framework for Assurance Engagements).

A doktori értekezés a következőképpen épül fel:

Az első részben kerül bemutatásra, miről is szólhat valójában egy ellenőrzés, milyen céljai lehetnek, milyen szakaszai vannak, milyen eszközök állnak rendelkezésre a célok eléréséhez, és végül, de nem utolsó sorban, hogy milyen környezeti feltételekkel szembesülünk egy ellenőrzés folyamán. Ahol lehetséges, kísérletet teszek majd arra, hogy a verbális jellegű célokat, eszközöket megfeleltessem a statisztika hasonló értelmű, szabatosabb fogalmainak, és ezt felhasználva rámutassak arra, hogy egyes, a gyakorlatban egyébként „bevett” eljárások alkalmazása csak a szükséges – és többnyire elég szigorú – feltevések fennállása esetén volna megengedhető.

A második rész az elemző eljárásokról szól. Ezek az eljárások napjainkban is domináns szerepet töltenek be a mindennapi alkalmazásokban, jelentőségük semmiképpen sem becsülhető le. Ezek közül különösen fontos szerepet tölt be a kockázatelemzés⁸, hiszen általában ez alapozza meg a legtöbb audit végrehajtási módját. Az elemző eljárások közül a kockázatelemzés általános ismertetése mellett külön is foglalkozom a számjegyelemzéssel, egy olyan új, de kissé „okkultnak tűnő” területtel, amely csak az elmúlt 10-15 évben kezdett

⁷ A megfelelő angol szakkifejezés: assurance engagement

⁸ A kockázatelemzést az ISA 315 szabvány 7. pontja alapján nem csak elemző eljárásokra kell alapozni, hanem emellett interjúkra, megbeszélésekre, és egyéb megfigyelésekre. Mivel terjedelmi okok miatt nem törekedhetek az ilyen értelemben vett teljességre, és ezek a módszerek egyébként is elég távol esnek a disszertáció fő csapásvonalától, ezért a kockázatelemzést mint az elemző eljárások egyikét kezelem.

elterjedni az ellenőrzési gyakorlatban. A külön tárgyalást indokolja, hogy a jövőben valószínűleg Magyarországon is egyre többet fogják használni ezt a módszert bizonyos gazdasági bűncselekmények felderítésében. A számjegyelemzés elméleti alapjául szolgáló Benford-törvénnyel kapcsolatosan megmutatom, hogy a törvény érvényesülése adott számrendszerben egyenértékű azzal, hogy az eloszlás karakterisztikus függvénye kielégít bizonyos, a számrendszer alapjától függő feltételeket. Ezt felhasználva bemutatok egyrészt egy új bizonyítást arra, hogy egyetlen eloszlás sem felelhet meg tökéletesen a Benford-törvénynek az összes lehetséges számrendszerben egyszerre, másrészt – mintegy az előbbi bizonyítás melléktermékeként – egy új magyarázatot arra, hogy a törvény ennek ellenére miért igaz mégis a gyakorlatban.

A harmadik rész foglalkozik a sokasági arány becslésére visszavezethető ellenőrzési eljárásokkal. Ilyenek például az úgynevezett megfeleléségi tesztek, de a később tárgyalandó értékösszeg becslési eljárások egy részének is a sokasági arány becslés szolgál elméleti alapul. Bemutatásra kerül a hagyományos és a bayesi módszerek mellett a szekvenciális eljárás is, amelyet eddig méltatlanul elhanyagolt az ellenőrzési szakma, ezzel kapcsolatosan fogom bemutatni az Állami Számvevőszéknél bevezetett új szekvenciális tesztelési eljárást és annak alkalmazási tapasztalatait. Az eljárások ismertetésén túl a harmadik rész tartalmaz egy elemzést is az aránybecslési módszerek megbízhatóságáról, melynek eredményei az aránybecslés széles körű felhasználhatósága miatt túlmutatnak a disszertáció keretein.

A negyedik rész próbálja „lefedni” a könyvvizsgálat, mint a legelterjedtebb audit tevékenység statisztikai szempontból érdemi részét, a sokasági értékösszeg becslést. Ez a rész kísérletet tesz egy új megközelítéssel közelebb jutni az egyes becslések valós megbízhatósági szintjének megítéléséhez.

Súlyos problémát jelentett az egész kutatás folyamán, hogy a terület extrém módon multidiszciplináris. Az ellenőrzés egy alkalmazott, és – jelenleg még – csak minimális mértékben formalizált, tehát inkább verbális diszciplína a statisztikához, és főleg a matematikához viszonyítva. Ezzel szemben rengeteg olyan nem triviális háttérismeretet, tapasztalatot igényel, amelyek – verbális jellegüktől függetlenül – nehéz olvasmánnyá tehetik az értekezést a kívülállók számára. Ugyanez fordítva is igaz: egy, az alkalmazásokban járatos auditor nehezen igazodik el a graduális képzésben oktatott meghaladó statisztikai témákban. Igyekeztem a lehetőségekhez képest közérthetővé tenni, és minél több példával is illusztrálni a mondanivalómat, de ahol már emelt szintű matematikát kellett alkalmazni, ott erre sokszor nem volt lehetőség. Szeretnék tehát előre is elnézést kérni a Tisztelt Olvasótól emiatt.

I. rész: Az ellenőrzés

Az ellenőrzés történelmi gyökerei, főbb céljai

A legutóbbi évtizedeket megelőzően az ellenőrzési szakma nagyon heterogén képet mutatott. A fő törésvonal évszázadokkal ezelőtt jelent meg, amikor is markánsan elvált egymástól a függőségi viszonyból fakadó ellenőrzés, és a független szervezet által végzett ellenőrzés. Noha mindkét fő típus célja valamiféle bizonyosság szerzése, az ellenőrzött folyamatokba való beágyazottságuk, az ellenőrzöttekhez való viszonyuk teljesen más. Ez abból a szempontból fontos, hogy a két típus két különböző megközelítést jelent, és ez még ma is tetten érhető bizonyos eljárásokban.

Az emberiség történelmének attól a pontjától, hogy kialakult valamilyen belső munkamegosztás, azonnal szükségessé vált az ellenőrzés is. A munkamegosztás, a szervező(őd)és célja és mozgatója mindig az volt és az lesz, hogy a kollektív cselekvés által a csoport nagyobb hasznosságot érhet el, mint nélküle. Ha a mikroökonómiai elmélet oldaláról akarjuk megközelíteni a témát, azonnal felmerül a kérdés, hogy ez a kollektív cselekvés, amely a csoportot az optimumba vezeti, vajon a csoport összes tagja számára is optimális-e. Amennyiben ugyanis nem az, akkor komoly készletétést érezhet minden „megrövidített” érintett az elvárttól eltérő viselkedésre, például a „potyázásra”.

Az egyéni és a kollektív szempontból optimálisnak tekinthető viselkedés eltérésén kívül más tényezők is veszélyeztetik a csoport közös céljának elérését: ilyenek lehetnek például a természeti katasztrófák, vagy más, rivális csoportok cselekedetei.⁹ A továbbiakban mindazokat a körülményeket, amelyek a szerveződés jellegéből (a tevékenységekből), vagy a külső környezet sajátosságaiból adódóan veszélyeztetik a szervezet céljait, *eredendő kockázatnak (Inherent Risk, IR)* nevezzük.

Konkrét ellenőrzési kontextusba helyezve, az eredendő kockázat *a hibás állítások képződésének elsődleges kockázata*, az ellenőrzött szervezet, rendszer, tevékenység olyan belső sajátossága, amely fogékonnyá teszi például lényeges hibák, szabálytalanságok képződésére, és amelyek valószínűen megtörténnek, ha a később definiálandó

⁹ A továbbiakban általában a szervezet szót fogom használni a fenti csoportosulások, szerveződések jelölésére.

kontrollrendszer nem előzi meg, nem tárja fel, illetve nem korrigálja azokat. Az ellenőrizendő állítástól függően számos körülmény, illetve tényező magas szintű eredendő kockázatot jelenthet, mint például: a tevékenységek összetettsége, bonyolultsága; nagyszámú érdekelt fél; nagy értékű eszközök, tranzakciók; földrajzilag szétszórt tevékenység; a tevékenységet végzők hiányos szakmai ismeretei, tapasztalatlansága; a vezetésre nehezedő nyomás; egyoldalú gazdasági vagy technikai függőség stb.

Egy szervezet eredendő kockázatai, ha a szervezet tevékenysége változatlan, általában csak hosszabb időszak elteltével csökkennek (ha egyáltalán). Jelentős változások, a feladatok módosulása, jogszabályi környezet változása, új rendszerek bevezetése esetén az eredendő kockázatok jelentősen megnőnek, s csak hosszabb idő elteltével csökkennek le újra egy normális, elfogadható szintre.

Hamar rájöttek a szervezetek vezetői, hogy az eredendő kockázatok többsége kezelhető: ha teljesen meg nem is szüntethetőek, de várható negatív hatásuk, vagy bekövetkezési valószínűségük csökkenthető. Azokat az intézkedéseket, eljárásokat, melyeket az eredendő kockázatok mérséklése érdekében foganatosítanak, *kontroll*oknak nevezzük.

A kontrollok kezdetben nem jelentettek mást, mint az erőszakot, vagy az erőszakkal való fenyegetést. Egy olyan társadalomban, ahol a tolvajokat általában elkapják, és annak, akit elkapnak, levágják a kezét, jelentősen csökken a lopás kockázata. Megfigyelhető, hogy a maffia jellegű szervezeteknél mennyivel hatékonyabb a kintlévőségek kezelése, mint egy „normális” szervezetnél. Nem véletlen ennek fényében, hogy egyre több, partnerkockázatnak erősen kitett, de jó hírű vállalkozás adja el rossz, társadalmilag elfogadott módon nehezen behajtható követeléseit kevésbé jó hírű „követeléskezelő vállalkozásoknak”.

A kontroll másik lehetséges módja abban rejlik, hogy a szervezet tagjait alaposan megválogatják. Amennyiben kisszámú, jó képességű, erős csoporttudattal rendelkező egyén alkotja a szervezetet, mint például a különféle szektáknál, karitatív szervezeteknél, akkor az eredendő kockázatok jelentős részét azáltal lehet kontrollálni, hogy a tagok önkéntesen a csoportérdeket helyezik előtérbe az önérdekkel szemben. Hasonló funkciót tölt be nagyobb szerveződéseknel az erkölcs, a morál, az etika, amelyek a szervezetbe való belépéstől (tipikusan a születéstől) fogva folyamatosan „beégetődnek” a szervezet tagjainak tudatába. Ha ez nem lenne, lehetetlen volna társadalomban élnünk, hiszen a pusztá önérdekkövetés teljesen szétzilálná a közösségeket.

Természetesen ez utóbbi összekapcsolható a fenyegetéssel is, például úgy, hogy nem megfelelő viselkedés esetén az érintettet „kizárják a klubból”, kiközösítik. Ha nem létezik érdemi alternatívája a szervezeten belüli létnek, ez a kiközösítéssel való fenyegetés megfelelő biztosítékot jelenthet a szervezetet belülről veszélyeztető eredendő kockázatokkal szemben. Jó példa erre még jelenleg is a „falu nyelve”, illetve a különféle szakmai önszabályozó testületek, kamarák.

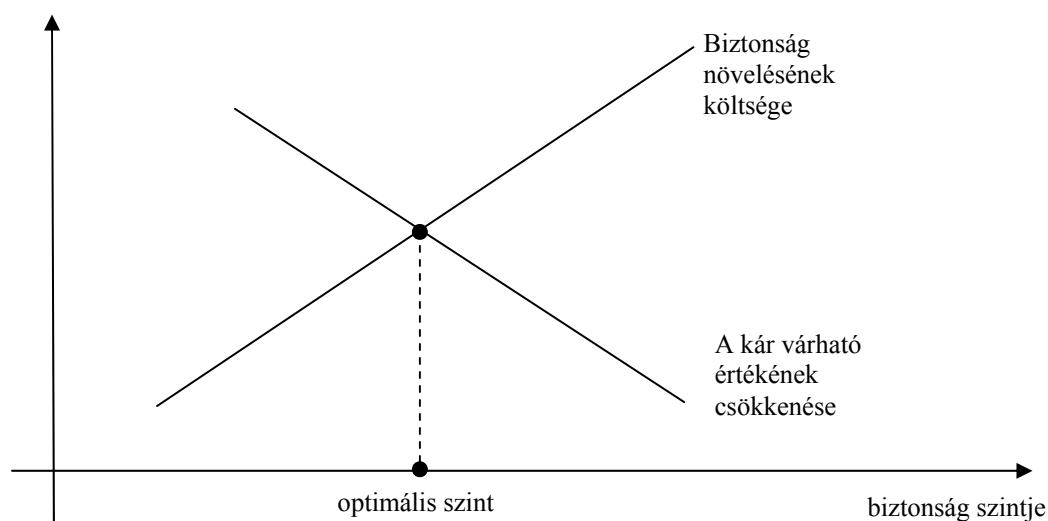
Külső, más szervezetek irányából származó kockázatok ellen is hasznos kontroll lehet a születéstől fogva belénk nevelt attitűd, mint például a hazaszeretet, az önfeláldozás stb. Ezek szigorúan nézve egyénileg nem feltétlenül racionálisak, mégis, a legtöbb ember gondolkodás nélkül szembemegy önérdékével, ha a helyzet úgy hozza.

Ha a szervezet mérete meghalad egy bizonyos szintet, az eredendő kockázatok kezelésére az előbbi két megközelítés (erőszak, önkéntes normakövetés) egyike sem jelenthet önmagában kielégítő választ. Az erőszakra épülő kontroll azért nem, mivel az ilyen kontrollok „optimális üzemmérete” korlátos, az önkéntes normakövetés költségei pedig a szervezet méretének növekedési üteménél gyorsabban emelkednek (nehéz nagyon sok jó embert összeválogatni).

A kapitalizmus egyik nagy vívmánya annak felismerése, hogy a kockázatok csökkenthetőek azáltal, ha a szervezetet meghatározott szabályok szerinti működésre készítetik. Ez tipikusan a munkafolyamatok pontos leírásában, és automatizmusok létrehozásában testesül meg. Míg korábban a ruhakészítéshez magas képzettségű, elhivatott szakemberekre volt szükség, a kapitalista világban a folyamatok átszervezése, automatizálása által képzettség nélküli betanított munkások is képesek voltak előállítani ugyanazt, vagy majdnem ugyanazt a végterméket. Ilyen módon alakult ki az a rendszer, amely jól átgondolt módon megszervezte a folyamatokat, automatizmusokkal biztosította az eredendő kockázatok mérséklését. Míg korábban a ruha minősége erősen függött a szabó aznapi hangulatától, és ez ellen csak a céh teljesen esetleges, önszabályozó mechanizmusai jelentették a kontrollt, addig egy gyárban az egymásra épülő kontrollok láncolata (elvileg) biztosítja a konstans minőséget, függetlenül az egyes textilipari munkások hangulatától.

Fontos azonban figyelembe venni, hogy a kontrollok működtetése nincs mindig ingyen: például a lopás elleni egyik lehetséges kontroll éjjeli őr alkalmazása, de szóba jöhet erősebb zárok, rácsok, vagy térfelügyelő kamerák felszerelése is, vagy ezek együttese. Köztudott, hogy ahová nagyon be akarnak törni, oda be is fognak, tehát teljes biztonság nem létezik. Mindazonáltal minél többet költünk biztonságra, annál kisebb lesz a lopás miatti veszteség valószínűsége.

Ha például a megvédendő raktárkészlet ellopásából 1 millió euró kárunk keletkezik, a lopás valószínűsége az adott időszakon belül 10%, a védekezés költsége pedig 1,5 millió euró, semmiképpen nem érdemes védekeznünk. Ezzel szemben 100 milliós lehetséges kárnál érdemes lehet megfontolni 1,5 millió euró befektetést is. A kontroll optimális szintjének meghatározása tehát a szokásos közgazdasági határelemzés keretében kezelhető. Ezt a megközelítést mutatja be az 1. ábra. (Vegyük észre, hogy ez a logika kiterjeszthető a lopástól eltérő kockázatokra is, mint ahogyan erre az ellenőrzés mélységének tervezésénél még utalni is fogunk.)



1. ábra: A lopási kockázat ellen működtetett kontrollok optimális szintje (sematikus ábra)

A fent leírt kontrollok a gyakorlatban nem feltétlenül működnek úgy, ahogyan elvileg kellene nekik. Hiába van például éjjeli őr, ha alszik az őrség alatt. Hiába van a portás, ha mindenkit beenged. És hiába van kettős aláírási kötelezettség előírva, ha nem tartják be. Ilyen esetekben nem tud megfelelő mértékben érvényesülni a tervezett hatás, nem mérséklődik eléggé az eredendő kockázatok bekövetkezési valószínűsége, várható hatása. Ezt a kockázatot hívjuk *kontroll kockázatnak* (**Control Risk, CR**). Jól érzékelteti a kontrollkockázat jelentőségét az a közismert jelenség, hogy az autózás sokkal veszélyesebb a repülésnél. Ennek nem az az oka, hogy a repülés eredendően biztonságosabb, hanem az, hogy ott sokkal szigorúbban betartják az előírt ellenőrzéseket.

A szervezetben alkalmazott kontrollok lehetnek manuális és automatizált kontrollok. A kettő típus között a fő különbség, hogy az automatikus kontroll „mindig” végrehajtódik, míg a manuális nem feltétlenül. Természetesen a számítógépes beépített kontrollokat szokás

példaképp hozni az automatikus kontrollra, azonban ezek nem a szabályzatnak, hanem a programkódnak „engedelmeskednek”, így véleményem szerint nem tökéletes példák. Sokkal jobb példa szerintem az automatikus kontrollra a felüljárók előtti magasságkorlátozó kapu: azon jármű tényleg csak úgy jut át, ha valóban alacsonyabb a kapunál. Elvileg az automatikus kontrollok olyanok, hogy ha egy esetben működnek, akkor mindig működnek. Az ilyen típusú kontrollok megvalósulásáról tehát teljesen máshogyan tudunk meggyőződni, mint egy manuális kontrollról. A disszertáció további részében (főleg az aránybecslési részben) ezért kontroll alatt általában manuális kontrollt értünk.

A manuális kontrollok potenciális áthághatóságára jelentenek megoldást az átfedő kontrollok: ezek olyan kontrollok, amelyek azt ellenőrzik, hogy a másik kontroll működött-e. Az átfedő kontrollok láncolata bármeddig fokozható, abszurd példaként akár az önkényuralmi rendszerekből ismert „mindenki figyel mindenkit” változatig is. Az átfedő kontrollok jelenthetik a kontroll teljes megismétlését (pl. ilyen a „négy szem elve”), de jelenthetik csak a kontroll-dokumentáltság meglétének ellenőrzését is (addig nem írja alá az átutalást a főnök, amíg nincsen rajta az „Ellenőrizve, fizethető” bélyegző). Ez utóbbi veszélyes, hiszen rutinból is lepecsételhetette az illető, valódi ellenőrzés nélkül.

Ellenőrzési kontextusba helyezve az eddigieket, a kontroll kockázat a belső irányítás, szabályozás (kontroll) kockázata, az a kockázat, hogy az ellenőrzött egység kontrollrendszere elmulaszt megelőzni vagy jelezni és kijavítani félrevezető állítást. A kontroll kockázatokat a szervezet aktív kockázatkezelési tevékenysége jelentősen csökkentheti.

A kontrollrendszerek kialakítása, működtetése meglehetősen termékeny táptalajt jelent a corporate governance irodalom számára. A világszerte széles körben elfogadott alapelveket jelenleg a COSO (Committee of Sponsoring Organizations of the Treadway¹⁰ Commission) nevű szervezet standardjai, „keretrendszerei” jelentik (Internal Control Framework illetve Enterprise Risk Management Framework)^{11,12}. A COSO a '70-es, '80-as években az USA-ban történt nagyvállalati botrányokat követő szabályozói aktivitás gyermeke. Számos kongresszusi meghallgatás történt és számos vizsgálóbizottság alakult akkoriban annak érdekében, hogy tisztázzák az „internal control” fogalom jelentését és meghatározzák a „jó”

¹⁰ James C. Treadway, a Securities and Exchange Commission elnöke vezette a bizottságot

¹¹ A COSO keretrendszereiről magyarázatokkal ellátott részletes ismertetés található például Root [2000] (Internal Control Framework) illetve Moeller [2007] (Enterprise Risk Management Framework) könyvében.

¹² Más kockázatkezelési rendszerek – például RIMS, CAF – közel sem annyira elterjedtek, vagy nem annyira univerzálisak, mint a COSO keretrendszerei, ezért terjedelmi okokból itt nem kerülnek ismertetésre.

kontrollrendszerek kötelezően elvárható minimális jellemzőit, ugyanis ebben látták a probléma megoldását. A Treadway bizottság jelentése 1987-ben jelent meg, amiben javasolták, hogy a menedzsment az éves beszámoló keretében tegyen jelentést az általuk működtetett belső kontrollrendszeréről is. A jelentés hangsúlyozta az erős kontrollkörnyezet, az etikai kódex, az audit bizottság és a hozzáértő menedzsment jelentőségét, továbbá annak fontosságát, hogy az „internal control” fogalmát széleskörű szakmai konszenzussal definiálják.

A Treadway bizottság működését finanszírozó szervezetek (ezek között domináns szerepet töltek be a könyvvizsgáló cégek, és az ennek betudható hangsúlyeltolódások miatt a COSO későbbi munkáit több kritika is érte – lásd pl. Root [2000]) a '87-es jelentést követően folytatták a munkát, és 1992-ben kiadták az ún. első keretrendszert (Internal Control Framework). Az első keretrendszer kiadását széles körben, több tízezer gyakorlati és elméleti szakember részvételével lefolytatott szakmai vita előzte meg az USA-ban, azonban ennek ellenére a keretrendszer tartalma és címe nincs összhangban: a keretrendszer a címnél sokkal szűkebb területre terjed ki, csupán a beszámoló készítésével kapcsolatos kontrollokkal foglalkozik (mintha csak az volna szervezetek legfőbb feladata, hogy éves beszámólót készítsenek). Később a keretrendszert többször kiegészítették, például a közszféra specialitásaival, de így sem tudta megakadályozni az évezred végén bekövetkezett újabb botrányokat, például a hírhedt Enron botrányt. A végtermék minden hiányosságától függetlenül azonban már magának a keretrendszernek a megalkotása is felbecsülhetetlen fontosságú, hiszen innentől fogva már volt mihez viszonyítani a könyvvizsgálati munka során.

Az első keretrendszer megjelenését követően, annak hiányosságait felfedezve kezdték el feldolgozni az internal controlnál jóval általánosabb fogalom, a risk management témakörét. Ahogy az internal control, úgy a risk management fogalmára sem volt konszenzusos definíció, ezt alkotta meg a második keretrendszer 2004-ben. A második keretrendszer már a teljes szervezeti működést átfogja, és ezáltal kiküszöböli az első keretrendszer legnagyobb hibáját, a beszámoló készítésre irányuló „csőlátást”. Mivel a keretrendszer megjelenése óta nem telt el sok idő, a világ még mindig csak „tanulja”, ezért a jövő fogja eldönteni, hogy mennyire fog bevélni. Tekintettel arra, hogy az túlságosan eltérne a disszertáció elején megfogalmazott céltól, a COSO keretrendszerek részletesebb ismertetésétől eltekintek.

Az USA-ban külső forrást (pl. részvény vagy kötvény formájában) bevonó szervezetek belső kontrollrendszerének kialakítását nem csak a COSO, vagy egyéb intézmények által

kidolgozott keretrendszerek befolyásolják, hanem – az Enron és hasonló botrányok következményeként – egy 2002-ben született, előterjesztőiről Sarbanes-Oxley Act-nek (röviden: SOx) nevezett törvény is befolyásolja. Ez a jogszabály több olyan kötelezettséget is előír, amelyek nemcsak a vállalkozásokat, hanem az auditor szakmát is érintik. A SOx többek között kötelezően előírja az első COSO keretrendszer használatát (korábban ez, akár a többi standard, nem volt kötelező érvényű), a menedzsereket személyesen teszi felelőssé a dokumentálásért, és ami a legfontosabb a jelen disszertáció szempontjából, a SOx 404-es szakasza kötelezően előírja a belső kontrollrendszer évenkénti felmérését és a felmérés eredményének nyilvánosságra hozatalát. Ez a rendelkezés tehát a jelen disszertáció III. részében leírt módszereket legalább egyenértékűvé teszi a IV. részben leírtakkal a gyakorlati fontosság szempontjából. A SOx részleteiről és a COSO ERM-mel való kapcsolatáról lásd Moeller [2007] könyvének 7. fejezetét.

Ha a szervezet mérete túllép egy szinten, a szervezet működése kaotikussá válhat, és leginkább feladat-delegálások láncolatával írható le. A feladat-delegálás annyit jelent, hogy a feladat elvégzésének felelőse nem saját maga végzi el a feladatot, hanem mást bíz meg vele (természetesen a felelősséget nem lehet delegálni).

A feladat-delegálás értelemszerűen együtt jár az elméleti közgazdaságtanban elég alaposan kielemezett megbízó-ügynök probléma kiteljesedésével. Ennek következtében a feladatot delegáló elemi érdeke, hogy a lehető legtöbb információnak jusson a birtokába, azaz, például, megkövetelje a folyamatok dokumentálását – ezt szokta a köznyelv (negatív jelentéstartalommal) bürokráciának hívni. Nem elég tehát elvégezni a feladatot, hanem a munkalapot is ki kell tölteni; azt pedig, hogy valaki ellenőrzött valamit, azt az aláírásával kell igazolnia. Hiába működhetne jól egy szervezet ezek hiányában is, a vezetés ezek nélkül nehezen tud megbizonyosodni arról, hogy a delegált feladatokat valóban elvégezték-e.

Sokszor azonban akkora a szervezet, hogy a dokumentálás sem elég: egyrészt a delegálónak nincs kapacitása arra, hogy ellenőrizze minden delegált feladat megfelelő elvégzését, másrészt a megbízó-ügynök probléma, az információs aszimmetria nem is teszi ezt mindig lehetővé. A dokumentáció nem jelent ugyanis teljes bizonyosságot, „a papír sok mindent elbír”. Ilyenkor a vezetők általában mást bíz meg az ellenőrzés elvégzésével: ők a belső ellenőrök (vagy a belső ellenőri funkcióban működő auditor/consulting cégek).

A szervezetek működésével kapcsolatos információra nem csak a vezetésnek van szüksége. A szervezet többi belső (pl. alkalmazottak) és külső (pl. állam, hitelezők, üzleti partnerek) érintette a vezetéshez hasonlóan igényel bizonyos információkat, melyek valóságtartalmáról igyekszik meggyőződni. Ennek megfelelően kialakultak a bizonyosság szerzésére szolgáló különféle eljárások, melyeket összefoglaló néven a továbbiakban auditnak, ellenőrzésnek¹³ fogunk nevezni. Az a szaktudás, amely lehetővé teszi a fenti értelemben vett bizonyosságszerzést, képessé teszi az auditorokat tanácsadásra is. Ez a szinergia azzal jár, hogy a gyakorlatban sokszor keverednek a tanácsadási (consulting) és a bizonyosságszerzési (auditing) elemek. A továbbiakban audit alatt csak a bizonyosságszerzést fogom érteni.

Fontos kis kitérőt tenni az ellenőrzés és a kontroll fogalmak kapcsolatának tisztázására. Amint a korábban leírtakból kitűnik, az arra jogosult személy a megszerzett információkat felhasználva intézkedik, „kontrollt gyakorol” a folyamatok felett: valójában ez nem más, mint menedzsment-funkció. Az információk megszerzésének egyik módja az ellenőrzés, tehát például egy ellenőrzés eredményét felhasználva kerülhet sor azután az intézkedésre. Az ellenőrzést végezheti közvetlenül maga az intézkedésre jogosult (például az átutalási megbízás aláírója ellenőrzi aláírás előtt, hogy helyes-e az összeg és a számlaszám), de végezheti az intézkedésre jogosult megbízottja is (például a belső ellenőr, akinek jelentése alapján a vezérigazgató intézkedik a folyamatok átszervezéséről). Attól függően, hogy az ellenőrzést végző rendelkezik-e intézkedési jogosultsággal is, nevezhetjük ugyanazt a tevékenységet kontrollnak, vagy ellenőrzésnek.

Az ellenőrzés-kontroll polémiaát szeretném egy példával is megvilágítani. A szervezeteket sok esetben nem közvetlenül emberek, sokkal inkább más szervezetek alkotják. A rész-szervezet belső ellenőrzési osztályának tevékenysége a nagy szervezet szempontjából már kontrollnak minősül, ugyanis az ő nézőpontjából ugyanaz a rész-szervezet (pl. mint jogi személy) végzi az ellenőrzést és az esetleges intézkedést is. A rész-szervezet szempontjából azonban ez továbbra is függetlenített belső ellenőrzésnek minősül, hiszen az ellenőrzést végző (természetes személy) és az intézkedést végző különböző. Összefoglalva tehát: *ugyanaz a tevékenység kontroll és ellenőrzés is lehet attól függően, hogy milyen „messziről” nézzük.*

Az ellenőrzés sok esetben elég barátságtalan formát is ölthet a gyakorlatban: például házkutatás, lehallgatás, és egyéb nyomozati tevékenységek. Noha ezek első ránézésre elég

¹³ A fentebb említett kontroll és a most bevezetett ellenőrzés nem csak a köznyelvben hasonló értelmű szavak. A két fogalom közötti kapcsolatot a következő bekezdés tisztázza.

ijesztőnek tűnnek, és főleg az állami érintettséggel tudjuk őket gondolatban összekötni, a valóság ennél komplexebb. Teljesen más tálalásban, de egy „normális” könyvvizsgálat során is majdnem ugyanez történik¹⁴. Akárcsak az adóellenőrnek, a „sima” könyvvizsgálónak is rendelkezésére kell bocsátani minden kért adatot, szabad bejárást kell neki biztosítani a helyiségekbe, a dolgozókkal személyes interjút (vö. kihallgatás) készíthet, hiszen ellenkező esetben nem fog tudni kellő bizonyosságot szerezni a témáról. A különbség mégis az, hogy az egyik esetben egy ellenállhatatlan külső kényszer (pl. tulajdonos, állam), a másik esetben az ellenőrzött saját döntése (pl. tőzsdére lépés) miatt kerül sor az auditra.

A felügyeleti szervek által elrendelt ellenőrzés tendenciájában sokkal inkább mutathat a potenciális felelősségre vonás irányába, mint a független szervezet által végzett ellenőrzés¹⁵. Egy hitelezőnek más adatok fontosak, mint egy részvényesnek. Egy vevő teljesen máshogy értékelhet egy termék-készletet, mint a biztosítótársaság, amelyik biztosítja. Ennek a sok érdekeltnek nehéz egyszerre megfelelni, a különféle ellenőrzési célkitűzések különböző módszerek alkalmazásával járnak.

Összességében azonban az ellenőrzési célkitűzések (objectives) a következő négy nagy csoport egyikébe besorolhatók¹⁶ (a zárójelben szereplő angol kifejezések az ellenőrzési cél nemzetközi szaknyelvben elterjedt megnevezései):

- valamilyen tétel *értékének becslése* ((accounting) estimation objective)¹⁷,
- *hitelesítés* (attestation objective)¹⁸,
- *valamilyen, jól definiált szabálynak való megfelelés ellenőrzése* (rule adherence, compliance, vagy propriety objective),
- *teljesítményellenőrzés* (value-for-money objective).

¹⁴ Természetesen lehetnek kivételek ez alól is, éppen úgy, mint ahogy rendőrségi kihallgatásra sem kerülhet mindig sor

¹⁵ Amennyiben bizonyos jogszabályi feltételek fennállnak (pl. pénzmosás gyanúja), a független auditorok is kötelesek bejelenteni az általuk találtakat, aminek a következménye rendőrségi nyomozás, bírósági ügy lehet.

¹⁶ Az ellenőrzési célok rendszerezésével kapcsolatosan lásd Leslie et al. [1980].

¹⁷ Ezt ugyan a legtöbb általam olvasott szerző nem sorolja az ellenőrzési célok közé (de például a financial audit technikájának egyik „alpműve”, Leslie et al. [1980] igen). A kihagyás mellett szól, hogy nem klasszikus bizonyossági szolgáltatásról van szó, sokkal inkább „kutatómunkáról”. Ennek ellenére – álláspontom szerint – jelen disszertációban érdemes önálló ellenőrzési célként kiemelni, hiszen a legtöbb ellenőrzés során előfordul ilyen jellegű feladat, a végrehajtandó művelet pedig – és a disszertáció szempontjából ez a lényeg – statisztikai nézőpontból teljesen más megítélés alá esik, mint a többi cél esetén (pontbecslés, és nem hipotézisvizsgálat).

¹⁸ Tipikusan a pénzügyi beszámolókhöz kötődő ellenőrzési cél: „financial statements are fairly presented”.

A filozófiai-történelmi okfejtés után visszatérünk az első bekezdésben leírt gondolatra, nevezetesen, hogy markánsan elvált egymástól a függőségi viszonyból fakadó ellenőrzés, és a független szervezet által végzett ellenőrzés. A hitelesítés, mint ellenőrzési cél történelmileg alapvetően a független auditorokhoz kötődik, míg a szabályoknak való megfelelés ellenőrzése sokkal inkább a felügyeleti ellenőrzés sajátja (vegyük például a szakhatóságok, mint tisztiorvosi szolgálat, közterület-felügyelet, stb. ellenőrzéseit). Jelenleg azonban már nem húzható ilyen egyértelmű határvonal a kettő közé. A magán- és a közszféra közeledése, a „New Public Management” filozófia terjedése, az átláthatóság iránti igény növekedése mind abba az irányba hat, hogy a közszféra beszámolóit is egy, a végrehajtástól független szerv hitelesítse. Ezzel párhuzamosan a magánszféra beszámoló hitelesítései nem szólhatnak csupán a hitelesítésről, az esetlegesen felmerülő szabályszerűségi (compliance) problémákat sem szabad figyelmen kívül hagyni, mivel azok visszahathatnak a beszámoló megbízhatóságára¹⁹. A teljesítményellenőrzési cél a legfiatalabb a négy célkitűzés közül, legfőbb alkalmazási területe a források hatékony, eredményes és gazdaságos hasznosításának megítélése, főként a közpénzek elköltésével összefüggésben.

Statisztikai szempontból – így a jelen értekezés szempontjából is – az adja az ellenőrzési célok csoportosításának relevanciáját, hogy ezek egy-egy jól körülírható (következtető) statisztikai feladatot képviselnek. A hozzá fűzött lábjegyzetből is kitűnik, hogy az első cél a pontbecsléssel azonosítható. A hitelesítés értékösszeg-becslési feladat. A szabályoknak való megfelelésről való megbizonyosodás általában a sokasági arány becslését jelenti. A teljesítmény ellenőrzése ebben a kontextusban egyebek mellett valamilyen komplex függvényformával megadott teljesítmény-indikátor függvény értékének tesztelését jelenti.

Egy ellenőrzésnek általában nem csak egy célja van: a négy fő ellenőrzési cél közül általában legalább 2-3 felmerül. Ennek megfelelően egy konkrét ellenőrzésben több technikát (és nem csak statisztikai technikákat, sőt, többségében nem statisztikai technikákat) is alkalmaznak párhuzamosan a megfelelő bizonyosság megszerzésére.

Fontos hangsúlyozni, hogy az *ellenőrzési célok* és a pénzügyi, gazdasági jellegű *ellenőrzések típusai nem esnek egybe*: az ellenőrzések tipizálásának alapja nem csak az ellenőrzés célja, hanem más is lehet, és emiatt az ellenőrzési típusok definiálása közel sem olyan egyértelmű, mint az ellenőrzési (rész)célok definiálása. Az INTOSAI csoportosítása alapján például két ellenőrzési típus létezik: „*audit of regularity*” és „*audit of management*”. Ezzel szemben az

¹⁹ például egy súlyosabb büntetés miatt

ISA standard által használt általános (tehát nem csak a könyvvizsgálatot – audit of financial statements – tartalmazó) ellenőrzés-fogalom („*assurance engagement*”, azaz kb. „bizonyossági szolgáltatás”) két lehetséges típusa a „*reasonable assurance engagement*” és a „*limited assurance engagement*”.

A leírtak miatt az ellenőrzés, mint szakma nem rendelkezik egzakt szabályokkal, „*a legjobb gyakorlat*” (best practice), vagy ha úgy tetszik, „*a jó gyakorlatok*” a nemzetközi és nemzeti szakmai szervezetek standardjaiban öltenek testet. Fontos szervezetek ezen a téren:

- az IFAC (International Federation of Accountants), amely többek között a financial audit típusú ellenőrzésekben is alapvető ISA (International Standards on Auditing) standardok folyamatos felülvizsgálatát végzi,
- az INTOSAI (International Organization of Supreme Audit Institutions), a közszféra külső ellenőrzési standardjaival
- az IIA (Institute of Internal Auditors), a belső ellenőrzés standardjaival.

Hazánkban a legfontosabb ilyen szakmai szervezet az Állami Számvevőszék és az Magyar Könyvvizsgálói Kamara.

Ugyan nem közvetlenül az ellenőrzéshez kapcsolódnak, de ezen a ponton kell szót ejtenünk a számvitelnek, mint a legfontosabb határterületnek a standardjairól. Történelmileg a számvitel országonként és alkalmazási területenként (magánszféra-közsféra) eltérően fejlődött, ám a globalizált világ kikényszeríti ezeknek a számviteli szabványoknak – GAAP (Generally Accepted Accounting Principles) – a közelítését. Ezeket tartalmazza az IFRS (International Financial Reporting Standards) és az IAS (International Accounting Standards), illetve az Egyesült Államokban alkalmazott US GAAP. Mivel ezek a standardok elvileg egységesen értelmezhetővé, összehasonlíthatóvá teszik a beszámolókat, a könyvvizsgálat sok esetben éppen azt tanúsítja, hogy az auditált beszámoló a nemzetközi standardok valamelyikének megfelelően készült el.

Az ellenőrzés folyamata

Az ellenőrzés elvileg a megbízással kezdődik, amikor is az auditorral ismertetik az ellenőrzés célját, céljait, megállapodnak az ütemezésben, stb. Ez a megbízás a valóságban nem feltétlenül formális, például a közsférában jellemzően jogszabályok definiálják az ellenőrzési

feladatot – jól vagy rosszul – és az ellenőrzést végző ezekhez igazodik, „visszakérdezésre”, tárgyalásra nincs sok lehetősége.

A megkezdett ellenőrzések első szakasza az ellenőrzött terület megismerésével, elemzésével telik. Az elemzés sokszor gyors és látványos eredményekhez vezet: csak az elemzés rámutathat olyan belső ellentmondásokra, amelyek magukban is „megalapozhatják” az elutasító véleményt. Ebben a szakaszban kerülnek többek között meghatározásra, azonosításra azok a részterületek is, ahol a lényeges hiba (hibák) nagy valószínűséggel előfordulhatnak: egy statisztikus szóhasználatával leírva ugyanezt, a lényeges hiba „prior eloszlását” próbáljuk így megadni.

Ez utóbbit hívjuk kockázatelemzésnek, ez az a terület, amiről a II. részben részletesen is szó lesz. A kockázatelemzés célja támpontot nyújtani az ellenőri erőforrások optimális allokálásához: minél kockázatosabb egy terület, annál nagyobb az esélye, hogy – amennyiben létezik – ott van a lényeges hiba.

Az audit általában úgy folytatódik, hogy azokat a kontrollokat, amelyeket az első fázis fontosnak minősített, és elvileg jónak talált, a gyakorlatban is tesztelik, ezáltal győződve meg arról, hogy tényleg működnek. Amennyiben ugyanis csak elméletileg, „papíron” működik egy kontroll, de a gyakorlatban nem, akkor nem tud érvényesülni a kockázatsökkentő hatása: meg kell húzni egy határt, ahonnan már nem tolerálható a kontroll-kimaradások aránya. A kontrolltesztek eredménye tehát további támpontul kell, hogy szolgáljon a kockázatok felderítéséhez. Ezzel a témával részletesen a III. rész foglalkozik, ahol is bemutatom a sokasági arány becslésére szolgáló módszereket.

Miután véglegesen meghatároztuk a kockázatos területeket, következnek az úgynevezett alapvető eljárások²⁰. Általában ezeket az eljárásokat egészen addig folytatják, amíg ésszerű mértékben sikerül meggyőződni pro vagy kontra az auditálandó állítás valóságtartalmáról.

Az alapvető eljárásoknak két típusát szokás megkülönböztetni: a részletek alapvető tesztelését, illetve az alapvető elemző eljárásokat. A részletek alapvető tesztelésére a legjobb példa, amikor konkrét adatok valóságtartalmát ellenőrizzük: ilyen lehet többek között, amikor meggyőződünk arról, hogy az elszámolt költségek valóban felmerültek-e. A részletek tesztelése az egyik legfőbb alkalmazási területe a statisztikai mintavételi és

²⁰ A vonatkozó angol szakkifejezés a substantive procedures. Az „alapvető eljárások”, vagy „alapvető vizsgálati eljárások” megnevezések meggyökeresedtek a hazai szaknyelvben.

következtetéseleméleti eljárásoknak. A könyvvizsgálatban tipikus értékösszeg becslési feladat szintén ebbe a kategóriába tartozik, erről bővebben a IV. részben fogok írni.

Az alapvető elemző eljárásoktól – a teszteléshez hasonlóan – bizonyítékot várnak a lényeges hiba meglétéről. Az adathalmaz belső struktúrájának elemzésével, a megelőző időszakok hasonló adataival való összehasonlítással sok esetben van erre lehetőség.

A mintavétel szerepe

Az ellenőrzés folyamatában több olyan szakasz is van, amikor nem feltétlenül kell az összes, rendelkezésre álló adatot feldolgoznunk. Ezekben az esetekben bőven elég, ha a teljes adathalmaz helyett annak csak egy részét vizsgáljuk meg, tehát mintát veszünk.

Tipikusan két olyan szituáció van az ellenőrzésben, amikor mintát veszünk erőforrás-takarékossági (legyen ez akár pénz, akár idő) megfontolásból:

- A minta alapján akarunk véleményt mondani a teljes sokaságról, azaz a mintából kiszűrt információt kivetítjük a sokaságra. Tipikusan ilyen, amikor valamilyen becslést készítünk. A mintavétel módját és a minta nagyságát ebben az esetben az határozza meg, hogy milyen típusú sokaságról milyen típusú véleményt akarunk mondani. El kell először is dönteni, hogy elégséges-e a sima pontbecslés, vagy pedig intervallumbecslésre, hipotézisvizsgálatra van szükség. Dönteni kell arról is, hogy a sokaság melyik jellemzője érdekel minket: pénzértékben vagy darabszámban (sokasági arányban) kifejezett választ várunk-e?
- Az auditor célja az, hogy a mintában lévő konkrétumok segítségével jobban megértse egy jelenséget, vagy nagy valószínűséggel megtaláljon bizonyos tételeket. *Egyik esetben sem fog levonni következtetést a teljes sokaságra nézve, noha mintát vett.* (Ezt én felderítő-elemző célú mintavételezésnek nevezem²¹, megkülönböztetve az előbbi, kivetítési célú mintavételtől. Az angol terminológiában egyébként a walk-through megnevezést szokás használni az elemző típusra.) Olyankor alkalmaznak az ellenőrök elemző célú mintavételt, amikor a szabályozás jobb megértése kedvéért gyakorlatban is végig akarnak követni egy folyamatot. Ilyen lehet például egy vaskos közbeszerzési

²¹ Ez nem keverendő össze a felderítési mintavétellel (discovery sampling), aminek célja gyors bizonyosság szerzése arról, hogy egy bizonyos jellegzetesség nem fordul elő a sokaságban.

eljárásrend megismerése konkrét közbeszerzéseken keresztül. Felderítő célzatú mintaválasztásra jó példa lehet a korrupciógyanús tételek megtalálása.

Látható, hogy a két helyzet két különböző mintavételi felfogást igényel. A korrektség kedvéért azonban meg kell jegyezni, hogy a valóságban ez a két dolog nem mindig válik el egymástól, és ez sokszor eléggé aggályos eljárásokat eredményez. Erre a legjobb példa az a bevett könyvvizsgálói gyakorlat, amikor célzottan kiválasztanak bizonyos, pl. nagy értékű tételeket, majd az ezekből levont következtetéseket kivetítik a sokaságra.

Alapesetben a kivetítési célhoz statisztikai mintát kell venni, a célzatos kiválasztás erre általában nem alkalmas. A felderítő-elemző célú mintavételezés pedig mind a statisztikai, mind a célzatosan kiválasztott mintával, mind ezek többlépcsős keverékével működik. A tapasztalatok szerint azonban a kivetítéshez is használható célzottan kiválasztott minta, ha az megfelelően reprezentálja a sokaságot, a felderítő-elemző célra pedig kevésbé hatékony egy egyszerű véletlen minta, mint a megfelelő más eszközökkel (pl. kockázatelemzés) kiválogatott célzott minta. Ez utóbbival összefüggésben elég csak arra gondolni, hogy a pénzügyi bűncselekmények előfordulási valószínűsége megnő ott, ahol a kontrollok gyengék, ezért ide fókuszálva a mintát várhatóan nagyobb „sikerrel” járunk. Hasonló ez ahhoz, ahogy a rendőrség sem véletlenszerűen kiválasztott embereket hallgat ki a nyomozás során, hanem célzottan, a saját szakmai szempontjai alapján hívja be az illetőket „interjúra”.

Statisztikai és nem statisztikai mintavételi eljárások

A mintavétel leírt két célja az előbb leírtak alapján általában más mintavételi megközelítést igényel. Ennek a két megközelítésnek a fő tulajdonságait tárgyalja ez a szakasz.

A statisztikai mintavétel az ellenőrzési standardok²² szerint olyan mintavétel, amely:

- véletlenszerűen választ mintát, illetve
- a minta kiértékeléséhez valószínűségelméleti eszközöket használ.

Amire ez nem igaz, az a nem-statisztikai mintavétel kategóriájába tartozik a standard értelmezésében. A statisztikai mintavétel egyik legfontosabb tulajdonsága, hogy számszerűsíthető a mintavételi kockázat: az a kockázat, hogy a mintavétel miatt az auditor

²² ISA 530. Az ellenőrzési standardoktól eltérően a statisztikai szaknyelv ezeket a fogalmakat máshogy használja: a statisztikai / nem statisztikai mintavételi eljárások helyett véletlen / nem véletlen eljárásokról beszél.

más következtetésre jut, mint amire a teljes sokaság vizsgálata alapján jutott volna (természetesen a nem mintavételi kockázatok, mint például a tételek felületes vizsgálata stb. itt is léteznek, de ezek nem számszerűsíthetők).

Az ellenőrzések során alkalmazott statisztikai mintavétel lehetséges módjai, mint például egyszerű véletlen, rétegzett, csoportos stb. általában megegyeznek az általános módszerekkel, azonban van néhány módszer, például a nem egyenlő valószínűséggel történő PPS mintavétel, amelynek a használata jóval hangsúlyosabb itt, mint más gyakorlati alkalmazásokban.

A könyvvizsgálat során a nem-statisztikai mintavétel főbb típusai az alábbiak²³:

- Nagy értékű, vagy kulcsfontosságú tételek (koncentrált kiválasztás). Az ellenőr valamely jellemző alapján választ ki tételeket, például azért, mert nagy értékűek, gyanúsak, szokatlanok, kockázatot rejtenek magukban, van már hibaelőzményük stb.
- Minden egyes tétel egy adott összeg felett. A könyvvizsgáló dönthet úgy, hogy megvizsgálja azokat a tételeket, amelyek értéke egy bizonyos összeget meghalad, és így ellenőrzi egy adott számlaegyenleg vagy tranzakció típus döntő részét. Ha a maradék tételek összértéke elhanyagolható, akkor ilyen módon kevés tétel megvizsgálásával is bizonyosság szerezhető arra nézve, hogy nincs lényeges hiba.
- Információt szolgáltató tételek, például konkrét folyamatokról, tranzakciókról stb.
- Eljárásokat ellenőrző tételek. Ezek a konkrét tételek alkalmasak lehetnek annak megállapítására, hogy egy adott eljárást alkalmaznak-e.

Az ellenőrzési gyakorlatban a mintavétel általában a statisztikai és nem statisztikai módszerek vegyes alkalmazását jelenti. A statisztikai mintavétel bizonyos módszerei, például a rétegzett és a csoportos mintavétel eleve igényelnek szubjektív előfeltevéseket. Az ellenőrzések során ehhez hozzájön még az is, hogy a statisztikai mintavétel alapsokasága általában egy nem-statisztikai minta.

Bizonyosság, pontosság, lényegesség

Hasonlóan a kontrollok működtetéséhez, az auditálás sincs ingyen. Minden egyes auditálással töltött órát magas óradíjon számláznak a cégek. Első megközelítésben persze azt lehetne

²³ A kategorizálás az ISA 530, 25 standardját követi.

gondolni, hogy az auditoroknak érdekükben áll minél „alaposabban” megvizsgálni a kérdéses területet, hiszen annál több lesz a bevételük, de ez nem igaz: versenyhátrányba kerülnének azokkal a cégekkel szemben, amelyek kevesebb munkaráfordítással tudják ugyanazt a bizonyosságot szolgáltatni.

Közkeletű, ám téves hiedelem, hogy amit az auditor megnézett és nem kifogásolt, az jó: kártérítési perek sokaságát indították már ilyen okból auditorokkal szemben arra hivatkozva, hogy például a később feltárt adóhiányt miért nem vette észre az auditáláskor. A valóságban azonban az auditor nem rendelkezik sem a bölcsék kövével, sem igazságszérummal, hasonlóan a bíróságokhoz, ahol néha előfordulnak téves döntések is. A bíróságokhoz való hasonlítás egyébként igenis helyénvaló: akárcsak a bíróság, az auditor is értékeli pro és kontra a bizonyítékokat, és ezek alapján adja meg a véleményét az adott témában. Ezen kívül a „téves” döntések itt sem mindig véletlenek (elég, ha az utóbbi évtized nagyvállalati botrányaira gondolunk, mint például Enron, Parmalat, vagy Xerox).

Az ellenőrzés során feltárt bizonyítékokkal kapcsolatosan a 2006. év végén hatályos ISA 500 szabvány úgy fogalmaz, hogy az auditornak elégséges mennyiségű és megfelelő minőségű bizonyítékot kell összegyűjtenie azon ésszerű következtetések levonásához, melyekre a véleményét alapozza²⁴.

A bizonyítékok elégséges mennyisége azt jelenti, hogy a bizonyítékok összességükben annyira meggyőzőek, hogy alátámaszthatják az auditori vélemény kiadását. Ez a definíció így fából vaskarika, valódi tartalommal csak a később definiálandó fogalmak ismeretében fog megtöltődni.

A bizonyítékok megfelelő minősége két szempontból értékelendő: a bizonyíték relevanciája és a bizonyíték hitelessége szempontjából.

Relevánsnak tekintjük a bizonyítékot, ha valamilyen módon hasznos lehet az ellenőrzési célnak megfelelő vélemény kialakításakor. Ha például szabályszerűségi ellenőrzést végzünk egy cégnél, akkor a kutatás-fejlesztési részleg elmúlt évi kiadásainak analitikus mélységű

²⁴ „The auditor should obtain sufficient appropriate audit evidence to be able to draw reasonable conclusions on which to base the audit opinion.” [ISA 500, 2] A fordítás során figyelembe vettem a hivatkozott szabvány 7. bekezdésének első mondatát, amely pontosítja az előbbi 2. bekezdést: „Sufficiency is the measure of the quantity of audit evidence. Appropriateness is the measure of the quality of audit evidence...”

ismerete általában nem sokat ér számunkra a szabályszerűség megítélésénél²⁵; igaz ez annak ellenére is, hogy például a konkurencia adott esetben sokat fizetne ezért az információért.

A bizonyítékok hitelessége sem egyforma. Általános hüvelykujjszabály az ellenőrzésben, hogy az interjúkon hallott információknál sokkal hitelesebb az ellenőr közvetlen tapasztalata; az írásban megerősített adatok hitelesebbek a szóbeli közlésnél; a harmadik, független féltől érkező adat hitelesebb a belső adatszolgáltatásnál (pl. banki egyenlegértékesítő – átutalási megbízások analitikája); a vezetőktől kapott tájékoztatás a beosztottaktól kapottnál. Ezek azonban csak hüvelykujjszabályok, a konkrét esetben a döntést az auditornak kell meghoznia.

A bizonyítékok utolsó, ám talán legfontosabb jellemzője a hozzáférhetőségük. A hozzáférhetőség is több dimenziós fogalom, hiszen a bizonyítékhoz való hozzáférésnek van időbeli (utánajárás, kutatás) és közvetlen pénzbeli (pl. levelezés, utazás) költsége.

A fenti öt tényező (elégségesség, hitelesség, relevancia, hozzáférés időigénye, hozzáférés költsége) együttes figyelembevételével kell döntenie arról, hogy milyen bizonyítékokat vizsgál, és milyen eljárásokkal az auditor.

Korábban szó volt róla, hogy az elégségesség definíciója fából vaskarika. A korábbi bekezdések segítségével azonban már képesek vagyunk azt tartalommal megtölteni.

Tegyük fel, hogy az auditor feladata a potenciális vevő megbízásából annak eldöntése, hogy a raktárkészlet értéke valóban 1 millió euró-e, mint ahogyan azt az eladó állítja. Tudni kell, hogy a raktárban lévő készlet fizikai mennyiségét valamilyen objektív ok miatt csak $\pm 3\%$ -os pontossággal lehet megállapítani (például, mert nem lehet a dobozok kinyitása nélkül ennél pontosabban megállapítani, mennyi a romlott árú). Ilyenkor természetesen a legjobb szándék ellenére sem lehet ± 30 ezer eurónál pontosabban becsülni a raktárkészlet értékét. Felmerül a kérdés, hogy vajon ez elegendő-e? Hajlandó-e kifizetni az 1 millió eurót a vevő, ha 3%-kal kevesebb árut kap? Ha igen, akkor ez a lehetséges 30 ezer eurós eltérés nem változtatja meg a döntését, amit máshogy úgy fejezhetünk ki, hogy az eltérés „nem lényeges”.

Lényegesnek nevezünk egy eltérést (hibát stb.), ha annak ismeretében más döntés születne, mint egyébként. Lényeges tehát az eltérés, ha befolyásolja a felhasználó döntését.

²⁵ Természetesen, ha a szabályszerűségi ellenőrzést a Gazdasági Minisztérium rendelte meg, hogy ezáltal győződjön meg a kutatás-fejlesztésre adott támogatás célnak megfelelő felhasználásáról, azonnal releváns az adathalmaz!

Hogy az adott audit során mi számít lényegesnek és mi nem, alapos megfontolás és/vagy konzultációk alapján kell eldönteni. Térjünk vissza ugyanis az előbbi példára. Mi van akkor, ha a 3% eltérés lényeges? Ilyenkor sajnos az eredeti eszközökkel nem lehet megfelelő pontosságú eredményt elérni²⁶. Amennyiben a megrendelő továbbra is ragaszkodik ahhoz, hogy adjunk neki véleményt a raktárkészlet értékéről, több dolgot is tehetünk:

- véleményt mondunk a meglévő információ alapján, ilyenkor azonban elég nagy a kockázata, hogy az esetleges pozitív véleményünk ellenére is lényeges eltérés volt az árban;
- kibontjuk egyenként az összes árut, tételenként megvizsgáljuk, és egy 1 500 000 eurós számla kíséretében átadjuk a készlet eurócentre pontos értékét;
- valamilyen módon mintát veszünk a sokaságból, kibontjuk, megvizsgáljuk, becslést készítünk a romlott termékek mennyiségére, és ez alapján adunk egy véleményt, amely esetleg tartalmazhat kisebb-nagyobb valószínűséggel kisebb-nagyobb pontatlanságokat.

Látható, hogy az első kettő rossz döntés. Az első esetben gyakorlatilag nem segítettünk semmit a megrendelőnknek, alig nyújtottunk neki bizonyosságot, hogy nem jár rosszul, ha megveszi a készletet. A második esetben tettük a lehető legrosszabbat: teljes bizonyosságot szereztünk ugyan a megrendelőnknek, de milyen áron! Többe került az auditori vélemény, mint amit egyáltalán a vásárlásra költött. Nyilvánvaló tehát, hogy a helyes út a harmadik. Meg kell találni a bizonyosságnak azt a szintjét, amely mellett a lényeges eltérések kockázata már elfogadható mértékű, a megállapítások kellően pontosak, és az egész nem igényel annyi munkát, hogy gazdaságtalanná tegye az auditot.

Ezt figyelembe véve tehát elégséges a bizonyíték, ha általa az elvárt mértékben biztosak lehetünk abban, hogy az audit tárgya nem tartalmaz lényeges hibát.

Mindezt a statisztikai nyelvére is lefordíthatjuk. A „fordítás” célja létrehozni egy olyan megfeleltetés-rendszert, aminek segítségével később könnyen elvégezhető az egyes ellenőrzési feladatok statisztikai vetületének pontos beazonosítása: semmiképpen nem cél

²⁶ Más példa ugyanerre a jelenségre: ha ezresre kerekített adatokkal dolgozunk, nem számíthatunk ezresnél pontosabb adatokra

tehát az ellenőrzési fogalmak újradefiniálása, a megfeleltetések alapvetően technikai célt szolgálnak.

- az audit az esetek többségében nem más, mint hipotézisvizsgálat: egy állítás (hipotézis) tarthatóságát kell eldönteni adott szignifikancia-szinten. Ritkán, de azért előfordul, hogy az audit célja nem hipotézisvizsgálat, hanem pontbecslés. Ez olyankor történik, ha a megbízó nem rendelkezik hipotézissel, igazolandó állítással, így magának az ellenőrzésnek a feladata ennek feltárása;
- a bizonyosság a statisztikához hasonlóan itt is a megbízhatósági- (vagy konfidencia-) szintet, tehát a szignifikancia-szint ellentettjét jelenti (ez pontbecslés esetén nem értelmezhető);
- a lényegesség pedig nem más, mint a hipotézisvizsgálat technikai nullhipotézise: azon lehetséges sokaságok „határa”, ahol elválik egymástól az elfogadható és a nem elfogadható (a lényegesség sem értelmezhető pontbecslés esetén). Amint látható, a lényegesség kérdése valójában mérési kérdés, meghatározásakor valójában egy absztrakt távolságfogalmat alkotunk a lehetséges sokaságok és az ideális sokaság között, és megadjuk a maximális, még elfogadható távolságot.

Könnyen lehet, hogy sokaknak ez az utolsó megfeleltetés meglepő, ezért külön is indoklom. Az első megfeleltetés alapján az audit általában egy hipotézisvizsgálat: van egy állítás, arról kell eldönteni, hogy igaz-e vagy sem. Tudjuk, hogy a hipotézisvizsgálat során két fajta hibát követhetünk el: elsőfajú és másodfajú hibát. Tudjuk továbbá, hogy ha az állítás a lényegesnél kisebb mértékben tér el a valóságtól, az ugyanolyan jó, mintha igaz lenne. Ha az állítás igazsága a nullhipotézis, akkor emiatt a nullhipotézisünk összetett: tartalmazza az állítást és a lényegesnél kevésbé eltérő állításokat. A technikai nullhipotézis a legkevésbé ellentmondó nullhipotézis, tehát az éppen lényeges hibát tartalmazó állítás. Ugyanezzel a logikával belátható, hogy ha az állítás hamisságát vesszük nullhipotézisnek, a technikai nullhipotézis ismét csak az éppen lényeges hibát tartalmazó állítás lesz.

Például: vegyük az előbbi esetet, azaz legyen az állítás az, hogy a raktárkészlet értéke 1 millió euró. Legyen lényeges az eltérés, ha a valós érték kisebb 970 ezer eurónál. Ez azt jelenti, hogy 970 ezer eurónál nagyobb raktárkészlet jónak minősül, ami kisebb, az rossznak. Bármilyen nullhipotézisnél ilyenkor a lényegesség, azaz a 970 ezer euró lesz a technikai nullhipotézis.

A nullhipotézis szerepe

Most, hogy rendelkezünk egy „szótárral”, elkezdhetjük alaposabb vizsgálat alá vetni az audit alapfogalmait. Az első kérdés, amire egy, a világ bármely pontjáról véletlenszerűen kiválasztott auditor 95% megbízhatósággal nem fog tudni helyesen válaszolni az, hogy mit jelent a 95%-os megbízhatóság, amiről olyan sokat beszélnek a prezentációs anyagokban.

Ugyan az előbbi mondat enyhe túlzás, valóban fontos dolog első kérdésként tisztázni, mit is jelent az auditori vélemény megbízhatósága. Ennek további vizsgálatához tekintsük a következő két táblázatot (1. Táblázat és 2. Táblázat), amelyek a bevezető statisztika könyvekhez hasonlóan a hipotézisvizsgálat során potenciálisan előforduló négy szituációt ábrázolják, először az audit, másodszor a statisztika nyelvén.

	Az állítás igaz	az állítás hamis
Pozitív vélemény	OK!	Mi mindent megtettünk.
negatív vélemény	Mi helyesen jártunk el. ²⁷	OK!

1. Táblázat: az ellenőrzés során potenciálisan előforduló négy (két) szituáció

	a hipotézis helyes	a hipotézis helytelen
elfogadó döntés	OK!	II. fajú hiba
elutasító döntés	I. fajú hiba	OK!

2. Táblázat: a hipotézisvizsgálat során potenciálisan előforduló négy szituáció

A két táblázat között nem egyértelmű a megfeleltetés: nem lehet tudni, hogy vajon mi legyen a nullhipotézis. Ismételten a bíróság a megfelelő példa: elvileg úgy is lefolytatható volna az eljárás, ha az ártatlanság véelme helyett a bűnösség véelmét használnák kiindulópontként. Tudjuk a statisztikából, hogy a másodfajú hiba sokkal kevésbé kontrollálható, mint az elsőfajú. Az is ismert, hogy a két lehetséges hiba közül az egyik az auditor, a másik az ellenőrzött (aki általában az ügyfél is egyben) kockázata. Nem meglepő tehát, ha az auditor a

²⁷ Figyelembe véve az ellenőrzési tevékenység bizonyosság nyújtó voltát, ez az álláspont (illetve a megfelelője a másik hibánál) teljesen helyénvaló (már ha tényleg igaz). A felhasználók azonban ezt nem nagyon tudják/akarják elfogadni, ami potenciálisan presztízsveszteséget, ezen keresztül később anyagi veszteséget is jelenthet az auditor cégeknek. Ennek a veszteségnek az elkerülése érdekében a cégek saját védelmükben óriási pótlólagos erőforrásokat használnak fel (és számláznak tovább) – pusztán ellenőrzés-szakmai okokkal – indokolhatatlanul.

saját kockázatát akarja kontrollálni (korlátok között tartani), ami a gyakorlatban annyit jelent, hogy nem szeretné a hibás állításokat igaznak elfogadni. Így tehát az audit 95%-os bizonyossága annyit jelent, hogy legfeljebb 5% valószínűsége van a téves elfogadásnak²⁸.

Az ellenőrzési kockázat jelentése, a „matematikai modell” kritikája

Az audit kockázat (az ellenőrzés kockázata) alatt – az előző résszel összhangban – statisztikai szempontból az első- és a másodfajú hiba elkövetési kockázatát értjük. Alaposabb vizsgálatot igényel azonban annak tisztázása, hogy pontosan miből is származik ez a kockázat.

Az ellenőrzés történeti bevezetőjében már definiáltuk a kockázat két típusát:

- az egyiket eredendő kockázatnak hívtuk, és úgy értelmeztük, mint mindazon tényezők összességét, melyek, megfelelő intézkedések hiányában veszélyeztethetik a szerveződés céljának elérését.
- a kontrollkockázat pedig azt jelenti, hogy az eredendő kockázatok kezelésére létrehozott kontrollok nem működnek, és ezért az eredendő kockázatok továbbra is érvényesülhetnek.

Amennyiben a szervezetet – leszűkítő értelemben – úgy tekintjük, mint aminek egyetlen célja az auditálandó állítás produkálása, akkor annak valószínűsége, hogy az auditálandó, de még nem auditált állítás lényeges hibát tartalmaz, kiszámítható a leszűkített szervezetre értelmezett eredendő (IR) és kontroll (CR) kockázatok valószínűségéből.

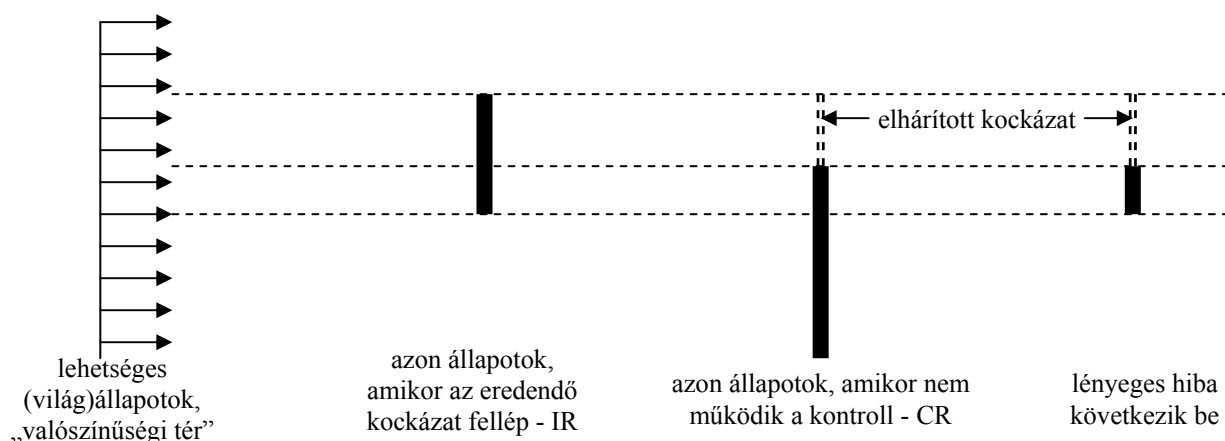
Ezt egy példán is demonstrálhatjuk. Tételezzük fel először, hogy ez a szervezet egyetlen eredendő kockázattal rendelkezik, és ennek elhárítására egyetlen kontrollt működtet. Valószínűségi számítási terminológiát használva ez két, egymást követő Bernoulli-kísérletet jelent: az első kísérlet során jelenik meg/nem jelenik meg az eredendő kockázat, a második kísérletben pedig működtetik/nem működtetik a kontrollt. Lényeges hiba akkor következik be, ha az eredendő kockázat megjelenik, és eközben a kontrollt nem működtetik. Amennyiben a két kísérlet független, a lényeges hiba bekövetkezési valószínűsége megegyezik a két komponens valószínűségének szorzatával, azaz formálisan

$$\Pr(IR \cap CR) = \Pr(IR) \cdot \Pr(CR) \quad (1.1)$$

²⁸ Ami egyben azt is jelenti, hogy a pozitív állításnak 95% a megbízhatósága (bizonyossága), ugyanis ha az állítás valójában hamis volna, akkor legalább 95% valószínűséggel az auditor ezt felderítette volna.

Amennyiben a két kísérlet nem független, csak alsó és felső határt tudunk mondani a valószínűségekre: az alsó határ természetesen 0, a felső határ pedig IR és CR valószínűsége közül a kisebb.

Érdemes „visszafordítani” auditor nyelvre ezt a két szélső lehetőséget: az együttes valószínűség akkor nulla, azaz akkor nem következik be soha lényeges hiba, ha a kontroll minden esetben működik, amikor kell. Ez egyébként az „ideális” kontroll jellemzője. Mivel mind a kontroll működtetése, mind a nem működése miatt bekövetkező esetleges hibák költséget jelentenek, a „legolcsóbb” megoldás esetén csak akkor működik a kontroll, ha szükség van rá. Ilyen kontrollra azonban nehéz konkrét példát felhozni. Talán ilyen példa lehet, amikor egy olyan épületet örököl a cég, amibe már beépítették az automatikus tűzoltó rendszert. Ha nincs tűz, a rendszer nem kerül semmibe; ha azonban tűz van, akkor automatikusan eloltja x költséggel. Ez a példa természetesen sántít, hiszen pl. a rendszer karbantartásával kapcsolatosan biztosan felmerülnek költségek, csak úgy, mint az érzékelők áramellátásával kapcsolatosan.



2. ábra: A lényeges hiba bekövetkezése

A másik szélső eset realisztikusabb: pont akkor nem működik a kontroll, amikor kellene. Az ellenőrzési tapasztalatokkal rendelkezők, de talán a laikusok is tisztában vannak vele, hogy sok gazdasági bűncselekményt olyan környezetben követnek el, ahol egyébként jól szabályozott, és általánosságban megbízhatóan működő kontrollok próbálnak gátat szabni az eredendő kockázatoknak, többek között a gazdasági bűncselekményeknek is. Mivel a tapasztalatok szerint az ilyen bűncselekményeket leginkább a felsővezetők követik el²⁹, az ő

²⁹ ISA 240

utasításukra a beosztottak áthágják a konkrét esetekben az előírt kontrollokat. De ha ennyire nem is nézünk extrém példát: az ISA standard szerint a csalásokat előidéző faktorok két nagy csoportja a motiváció és a lehetőség megléte. Amennyiben „lyukas” a kontrollrendszer, sokan, akik egy jó kontrollrendszer mellett nem mernének lépni, a lehetőséget látva csalást fognak elkövetni. Mindezeknek az a következménye, hogy sok esetben a kontrollok nem működése és az eredendő kockázatok megjelenése együtt járnak és nem függetlenek.

A kezdeti feltételezésen enyhítve, most engedjük meg, hogy több eredendő kockázat, és azok ellen több kontroll is létezzen. Legyenek tehát az eredendő kockázatok IR_1, IR_2 stb., és a kapcsolódó kontrollkockázatok az i . eredendő kockázat elleni kontrolloknál CR_{i1}, CR_{i2} stb. Annak valószínűsége, hogy lényeges hiba történik egy ilyen rendszerben:

$$\Pr\left(\bigcup_i \left(IR_i \cap \bigcup_j CR_{ij} \right)\right). \quad (1.2)$$

Ha az egyes eredendő kockázatok egymáshoz viszonyított struktúrája nem ismert – és a valóságban általában ez a helyzet – nem sokat tudunk mondani ennek a valószínűségnek az értékéről.

Az elmúlt bekezdésekben leírt gondolatmenet érzékeltette, hogy ugyan az eredendő- és a kontrollkockázatok bekövetkezési valószínűségének ismerete szükséges a lényegesen hibás állítás előfordulási valószínűségének megállapításához, azonban korántsem elégséges. Szükséges még ismerni az egyes kockázatok egymáshoz való viszonyát, ennek ismeretében számolható csak ki a nem auditált állításoknál a lényeges hiba valószínűsége.

Az (1.2) képlet helyett az ellenőrzési irodalomban szinte kizárólag az (1.1) képletet használják azzal a változtatással, hogy a Pr valószínűség operátort nem teszik ki. Az $IR \times CR$ formula önmagában teljesen korrekt lehetne abban az esetben, ha a kockázatokat, mint eseményeket értelmezzük (ilyenkor a szorzás művelet egyenértékű a halmazok közötti metszettel), ám amint a kockázatok valószínűségét értjük alatta, azonnal ingoványos talajra lépünk.

Elvileg feltételezhetnénk, hogy az IR és CR kockázatok képviselik az összes IR_i és CR_{ij} kockázat egyesítését, de könnyen látható, hogy ez az értelmezés téves eredményre vezet: nem

áll fenn ugyanis az $\bigcup_i \left(IR_i \cap \bigcup_j CR_{ij} \right) = \left(\bigcup_i IR_i \right) \cap \left(\bigcup_{i,j} CR_{ij} \right)$ azonosság. Ha csak egy

kockázatot és egy kontrollt feltételezünk a szervezetnél, akkor az $IR \times CR$ formula igaz lesz

arra az alesetre, amikor a két kockázat független. A függetlenségi feltételezés azonban – a korábban említett okok miatt – nagyon korlátozónak tekinthető.

Kijelenthető tehát, hogy az ellenőrzési szakirodalomban leírt $IR \times CR$ formula önmagában nem, vagy csak erős megszorításokkal alkalmas a nem auditált állításban lévő lényeges hiba valószínűségének meghatározására. Ennek ellenére a formula hasznosnak tekinthető, hiszen matematikailag nem helytálló ugyan³⁰, de jelzi a két kockázatnak a lényeges hibával való összefüggését. Egyáltalán nem helyénvaló azonban, hogy a formulát sokszor számolásra is felhasználják³¹.

Alaposan kielemezve a nem auditált állításban lévő hiba kockázatát, amit a szokásoknak megfelelően – elfogadva a formulában lévő hibákat – ezentúl én is $IR \times CR$ -rel fogok *jelölni*, itt az ideje, hogy meghatározzuk az auditált állításban lévő hiba kockázatát. Az *auditált állításban lévő lényeges hiba kockázata* (Audit Risk, AR) – röviden audit kockázat vagy ellenőrzési kockázat – azt a kockázatot jelenti, hogy a pozitív auditori vélemény ellenére az állítás lényeges hibát tartalmaz. Ez akkor fordul elő, ha

1. valamely eredendő kockázatból kifolyólag megjelenik a lényeges hiba,
2. ezt a létező vagy nem létező kontrollok láncolata az auditálást megelőzően nem hárítja el és végül
3. az auditor sem deríti fel.

Az utóbbi kockázatot *feltárási kockázatnak* (Detection Risk, DR) hívjuk. A feltárási kockázatnak a nem auditált állításban lévő hiba meglététől való függetlenségét jóval könnyebb intuitíve elfogadni, mint az eredendő és kontrollkockázat függetlenségét (noha nem kizárható, hogy az auditor összejátszik az ellenőrzöttel). Ezt szem előtt tartva felírható (legalábbis formálisan) az auditált állítás lényegesen hibás voltának kockázatára a következő képlet, melyet az auditorok az ellenőrzési kockázat matematikai modelljének neveznek.

$$AR = IR \times CR \times DR \quad (1.3)$$

Az (1.3) képlet sokszor átrendezik³², és kifejezik belőle a feltárási kockázatot:

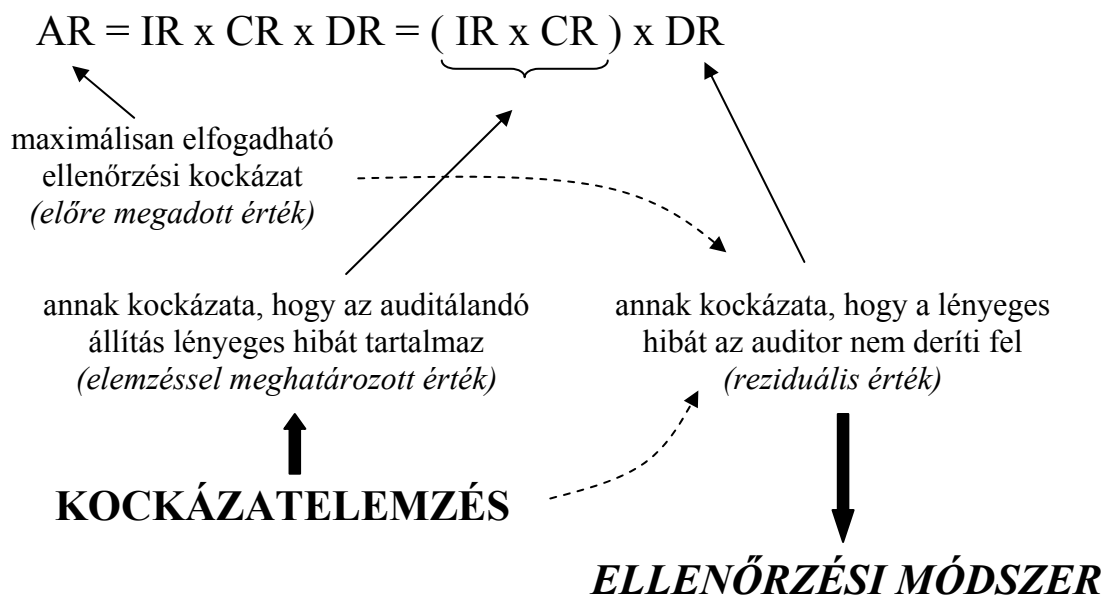
³⁰ A formula elméletileg nem helytálló volta több kutatónak is feltűnt, ez irányban történtek is kutatások (lásd például Cushing et al. [1983], Dhar et al. [1987], Waller [1993] vagy Haskins et al. [1995]), de – sajnálatos módon – különösebb gyakorlati visszhang nélkül.

³¹ Mint például az Állami Számvevőszék „zöld könyvében” (Lévai [2004]) is, a szükséges feltárási bizonyosság számszerű meghatározásánál (lásd még az ellenőrzési kockázat matematikai modelljénél leírtakat)

³² A képlet átrendezése akkor indokolt, ha a szorzásjelek értelme szorzás, és nem halmaz metszet. Ebben az esetben az IR, CR, DR, AR nem a kockázatok, hanem azok valószínűségét jelentik. Továbbra is fennáll tehát,

$$DR = \frac{AR}{IR \times CR} \quad (1.4)$$

Hogyan kell olvasni ezt a képletet? Az auditorok ebből a képletből szokták „meghatározni”, mennyit kell „dolgozni”. A kisebb feltárási kockázat elérése a nagyobb minta miatt több munkát jelent: kisebb esélyt adunk annak, hogy a nem auditált állításban lévő, lényeges hibát nem vesszük észre, ez pedig további bizonyítékok gyűjtését és elemzését igényli. Az (1.4) formula azt mondja ki, hogy a szükséges erőfeszítés mértéke az auditált állítás elvárt megbízhatóságának növekvő függvénye, ezzel szemben csökkenő függvénye a kontrollok megbízhatóságának. Ez utóbbi különösen fontossá teszi az auditor számára, hogy felmérje a releváns kontrollok megbízhatóságát az ellenőrzött szervezetnél.



3. ábra: Az ellenőrzés kockázati tényezőinek kapcsolata a választott ellenőrzési módszerrel

A feltárási kockázat (csakúgy, mint az eredendő és a kontrollkockázat) az esetek többségében nehezen számszerűsíthető. A később ismertetendő eljárások közül gyakorlatilag csak a statisztikai mintavételen alapuló következtetések olyanok, hogy a nem-feltárási valószínűsége pontosan meghatározható. Ez összességében azt jelenti, hogy az audit kockázati tényezőinek

hogy a képlet csak IR és CR (mint kockázatok, azaz események) függetlensége mellett igaz, számolni csak ekkor szabad vele.

közelítő számszerűsítése, és így az optimális munkaráfordítás meghatározása is sokkal könnyebben kivitelezhető a statisztikai eljárások használata esetén.

II. rész: Elemző eljárások

Az elemző eljárások célja és főbb jellegzetességei

Az „elemző eljárások” gyűjtőfogalom mindazon módszerek összefoglaló megnevezése, amelyek az adathalmazban való hibát az adathalmaz külső adatokkal való inkonzisztenciájának, vagy belső ellentmondásainak segítségével mutatják ki. Ez gyakorlatilag nem jelent mást, mint a köznap értelemben vett „józan ész”. A legtöbb ember nem fogja elhinni egy 8 lábú, pókszerű és pókméretű élőlényről, hogy az sasmadár; hasonlóképpen, egy könyvvizsgáló is nehezen fogja elhinni, hogy nagyjából változatlan külső és belső termelési feltételek esetén az előző évhez képest 93-szorosára növekedett a cég egy főre jutó termelése.

Bizonyos elemző eljárások eredménye egyértelmű elutasítást von maga után, más esetben csak az egészséges kételkedés amúgy is meglévő parazsát szítja fel. Már az is gyanús lehet az előző példában, ha a termelékenység csak 1,5-szeresére nőtt volna 93 helyett, de azt, hogy az adat hibás, nem lehet annyira egyértelműen kijelenteni, mint a 93-szoros esetben. Az elemző eljárások egyik alkalmazási területe ennek megfelelően a kockázatelemzés, ahol is azon részterületeket jelöljük ki segítségével, amelyek nagyobb eséllyel tartalmaznak hibát.

A könyvvizsgálatban az elemző eljárások során a beszámolóban lévő pénzügyi adatokat például a következőkkel szokás például összevetni:

- előző időszakok hasonló adatai;
- korábbi előrejelzések;
- a hasonló területen tevékenykedő egyéb cégek megfelelő adatai.

Ezen kívül érdekes tanulságok vonhatóak le az adathalmaz belső elemzéséből is, mint például a különböző területek haszonkulcsainak, vagy a bértömegnek és dolgozók számának összehasonlításából.

Az elemző eljárások lefolytatására sok eszköz használatos, kezdve az előbbi példától a komplex statisztikai elemzésekig. Tekintettel azonban arra, hogy a gyakorlatban használt elemző eljárások többsége inkább számviteli, vállalatértékelési stb. összefüggéseken alapul,

ezért – a dolgozat terjedelmi kereteit is figyelembe véve – csak azon elemző eljárások kerülnek kiemelésre, amelyek valamilyen statisztikai módszeren alapulnak, és az ellenőrzés során való alkalmazásuk érdemben eltér a többi terület hasonló alkalmazásától. Többek között ezért maradt el például a teljesítményellenőrzéshez kapcsolódó elégedettségmérések statisztikai eljárásainak feldolgozása is, ugyanis az valójában a közvélemény-kutatásoknál használt módszerek közvetlen alkalmazását jelenti.

Az elemző eljárások alkalmazása alapvető eljárásként

Noha a bizonyosság megszerzésének „közismertebb” módja a tesztelés, az ellenőrzés során ez sokszor kiegészíthető, esetleg ki is váltható elemző eljárásokkal. A hazai könyvvizsgálói gyakorlatban (és a könyvvizsgálók oktatásában is) – tapasztalataim szerint – az elemző eljárások súlya többszörösen meghaladja a mintavételes tesztelését.

Az elemző eljárások alkalmazásáról, illetve az alkalmazás mértékéről, területeiről való döntés során több tényezőt is figyelembe kell venni³³.

Az alapvető elemző eljárások tipikus alkalmazási területét jelentik a *kiszámítható, megjósolható tömegjelenségek*. Az alkalmazásnak tehát mindig azon az implicit feltevésen kell alapulnia, hogy az adatok között összefüggés van és lesz, hacsak nem állnak fenn olyan ismert körülmények, amelyek módosítják ezt az összefüggést (például egy új technológiát vezettek be, ami teljesen megváltoztatta a termelés módját). Az ilyen összefüggések fennállása bizonyítékként szolgálhat az ellenőrzött adatok teljességével, pontosságával és bekövetkezésével kapcsolatosan. Annak megítéléséhez azonban, hogy mennyire lehet egy adott elemzés eredményére támaszkodni a véleményalkotás során, szükséges felmérni annak a kockázatát is, hogy a lényeges hiba megléte ellenére is fennállnak a várt összefüggések, azaz nem inkonzisztens az adathalmaz.

A másik figyelembe veendő tényező *az elemzett adatok megbízhatósága*. Ezt befolyásolhatja azok forrása és jellege, illetve azok a körülmények is, amelyek között megszerezték az adatokat. Annak megállapításánál, hogy az adatok megbízhatóak-e az alapvető elemző eljárások megtervezése szempontjából, a következő szempontok léteznek:

³³ ISA 520, 12

- A rendelkezésre álló információ forrása. Az információ rendszerint megbízhatóbbnak minősül, ha az ellenőrzött szervezeten kívüli, attól független forrásból szerezték be. Egy banki egyenlegközlés például megbízhatóbb, mint a szervezet saját nyilvántartása.
- A rendelkezésre álló információ összehasonlíthatósága. Például, előfordulhat, hogy az általánosabb iparági adatokat ki kell egészíteni ahhoz, hogy összehasonlíthatóak legyenek a speciális termékeket termelő és értékesítő gazdálkodó adataival.
- A rendelkezésre álló információ jellege és relevanciája. Például, az előrejelzéseket inkább prognosztizált eredményeknek, vagy elérendő céloknak tekintik-e.
- Az információ előállításának módja, például, az előrejelzések előállításának, áttekintésének és betartásának az ellenőrzése.

Az elemző eljárásokat a fentiek miatt természetes módon kiegészíti azon kontrollok tesztelése, amelyek az elemzett információk előállításával kapcsolatosak. Ha ezek hatékonyak, akkor a generált adatok, illetve az adatokon alapuló elemzések eredményei is megbízhatóbbak.

Az elemzésektől elvárható pontosság mértéke, és ezzel párhuzamosan az eredmények megbízhatósága erősen függ az alapadatok pontosságától/stabilitásától. Ezt több tényező is befolyásolja:

- Mennyire várható el stabil viselkedés az adatoktól, mennyire pontosan lehet megjósolni a viselkedésüket. Például, a könyvvizsgáló rendszerint nagyobb következetességet vár az egyik időszakhoz kapcsolódó bruttó nyereségnek egy másik időszaki bruttó nyereséggel való összehasonlításánál, mint az olyan önként vállalt ráfordítások összevetésénél, mint a kutatás vagy reklám.
- Az adatok milyen mértékig bonthatóak meg, mennyire részletesek, és nem utolsósorban pontosak. Milliós értékre kerekített indikatív adatokból nem lehet forintra pontos eredményt kapni.

A korábban definiált lényegességhez kapcsolódik az elemzésektől *elvárt pontosság* fogalma, mint az utolsó olyan tényező, amelyet az elemző eljárás alkalmazásáról való döntéskor figyelembe kell venni. Az elemzéstől elvárható és elvárt pontosság viszonya jól mutatja, hogy az adott elemző eljárás által milyen mértékű bizonyosság szerezhető.

Az elemző eljárások alkalmazása a kockázatelemzésben

A szervezetnek és környezetének jobb megértésére, a kockázatok felmérésére az auditornak elemző eljárásokat kell használnia. Ezeket az elemzéseket mind a pénzügyi, mind a nem pénzügyi információkra külön-külön, mind ezek összhangjára nézve érdemes lefolytatni.

A kockázatelemzés célja valójában nem más, mint az ellenőrzési erőforrások optimális allokálása. Mivel az ellenőrzés célja adott megbízhatósággal véleményt mondani egy állításról, és az ellenőrzés nullhipotézise az állítás hamissága, ezért mindig olyan ellenőrzési módszert kell kiválasztani, amely mellett egyrészt a lényeges hiba felfedezési valószínűsége legalább a megbízhatósági szint, másrészt pedig ehhez minimális erőforrást használ. Ezt az optimumot úgy lehet elérni, ha adott ráfordítás mellett maximalizáljuk a lényeges hiba felderítésének valószínűségét.

Ez elméletben egyszerűnek tűnik, de valójában nem is olyan könnyű megtalálni a fenti optimumot. Magának az optimumkeresésnek a neve kockázatelemzés, lefolytatására több eljárás is ismert, ezeket adott esetben párhuzamosan kell alkalmazni. A lehetséges eljárások egyik nagy csoportját az elemző eljárások alkotják, a többi eljárással³⁴ terjedelmi okokból disszertációmban nem foglalkozom.

A kockázat dimenziói, mérése

A kockázat kétdimenziós fogalom, ellenőrzési kontextusban azt mutatja, hogy az egyes részterületek milyen valószínűséggel és mekkora mértékben járulnak hozzá az ellenőrzött állításban esetlegesen meglévő lényeges hibához.

Az egyik dimenzió, amellyel a köznyelv³⁵ – sokszor, és hibásan – azonosítja a kockázatot, a bekövetkezési valószínűség, a másik dimenzió pedig a bekövetkezéskor kiváltott hatás, amit a kockázat jelentőségének is hívnak. Valójában tehát a kockázat hasonlatos egy (többnyire karakterisztikus) valószínűségi változóhoz, azonban nem feltétlenül valós számokat (0/1-et) vesz fel értéként. Sokszor ezért hasznosabb a kockázatot úgy tekinteni, mint egy olyan

³⁴ Ilyenek például az interjúk a szervezeten belüli és kívüli személyekkel, közvetlen megfigyelések, stb.

³⁵ Sajnos a köznyelven kívül más is, például az államháztartás működési rendjéről szóló 217/1998. (XII. 30.) Korm. rendelet, amely így határozza meg a kockázat fogalmát: „63. kockázat: a költségvetési szerv gazdálkodása tekintetében mindazon elemek és események **bekövetkeztenek a valószínűsége**, amelyek hátrányosan érinthetik a szerv működését” [kiemelés tőlem, LT]

(valószínűségi számítási értelemben vett) eseményt³⁶, aminek jól definiált hatása („következménye”) van. (Ez utóbbit az ellenőrzési kontextusban természetesen szintén az ellenőrzött állítással összefüggésben kell értelmezni.) Nem bonyolítom feleslegesen a tárgyalást azzal a lehetőséggel, hogy a kiváltott hatás nem konstans, ugyanis ezt a lehetőséget egyetlen általam ismert ellenőrzési módszer sem veszi figyelembe. A következő bekezdésekből ki fog tűnni, hogy a kockázat mérése erősen szubjektív, több ok miatt sem lehet precíz, tehát a kockázatfogalom ilyen típusú kiterjesztésének (nem konstans hatás) az ellenőrzésben nem volna semmilyen gyakorlati haszna³⁷.

Annak érdekében, hogy a kockázatok hatásait figyelembe tudjuk venni, azokat valahogyan mérnünk kell tudni. A statisztikában definiált négy mérési szint – névleges, sorrendi, intervallum és arány – közül legalább sorrendi skála szükséges ahhoz, hogy összehasonlíthassunk két kategóriát. A kockázatok hatásai jó esetben már eleve arány skálán mérhetőek, valós számmal is megadhatóak: ilyen például az a kockázat, amelyet egyes termékek piaci árának megváltozása jelent más termékek árára, hiszen ilyenkor pénzben mérhető a hatás. Más esetekben képezhető intervallum, vagy arány skála az egyes hatások „hasznosságán” keresztül, hasznossági függvény által. Végül, sok esetben nem adható meg más, csak egy sorrendi skála, mint például alacsony-közepes-magas, esetleg 1-től 5-ig, vagy 1-től 10-ig pontozás.

A hatások ismerete, pontosabban azok mérése azonban nem elegendő ahhoz, hogy az egyes kockázatokat értékelni tudjuk. Noha egyszerűnek tűnhet, nem csekély módszertani problémát jelent az első dimenzió, azaz a bekövetkezési valószínűség meghatározása is.

Egyes esetekben, például egy tőzsdei portfólió esetén megadható „pontos” valószínűség bizonyos események bekövetkezésére. Természetesen ez a valószínűség, mint a gyakorlatban használt szinte minden valószínűség, szubjektív: függ attól az információs halmaztól, amivel éppen rendelkezünk: az, aki bennfentes információkkal rendelkezik, teljesen máshogy ítéli meg egy bizonyos árfolyam-elmozdulás valószínűségét, mint az, akinek csak a publikált információk állnak rendelkezésére. Pont ez a szubjektivitás az, ami miatt nem lehetnek nagyon szigorú elvárásaink a kockázatok bekövetkezési valószínűségének megítélésekor.

³⁶ Tehát a lehetséges kimenetek egy részhalmazát.

³⁷ Ez is egy példa arra, hogy ugyanazzal a névvel nevezett dolgok, ez esetben a kockázat fogalma, mennyire más jelentést nyernek egyes alkalmazásokban. Az említett kiterjesztés ugyanis teljesen helyénvaló egy értékpapír kockázatának elemzésekor, ahol is ténylegesen egy valószínűségi változóról, illetve annak eloszlásáról van szó tehát a „hatás” nem konstans.

Olyan események bekövetkezési valószínűségét, amelyek a múltban már többször bekövetkeztek, és valamilyen rendszersajátosság miatt állandó mintázatot követnek, a nagy számok törvénye miatt könnyen becsülhetjük a múltbeli bekövetkezési aránnyal. Ez a módszer azonban nem működik akkor, ha az esemény bekövetkezése közvetlenül nem figyelhető meg, vagy pedig az esemény olyan ritka, hogy nem áll rendelkezésre elegendő megfigyelés az empirikus valószínűség kiszámításához. Probléma léphet fel akkor is, ha nincs, vagy nem ismert olyan rendszersajátosság, ami biztosítaná a relatív gyakoriság stabilitását.

Példaként tekinthetjük a vezetők által elkövetett gazdasági bűncselekményeket. Ismert, hogy több tényező is befolyásolja annak valószínűségét, hogy elkövetésre kerül-e gazdasági bűncselekmény:

- léteznek-e ösztönzők, vagy más hasonló nyomás (akár szervezeten belülről, például prémiumcélok formájában, akár kívülről, például szerencsejáték-függőség),
- van-e kedvező lehetőség (például szabad hozzáférés a bankszámlához),
- milyen az illető morális hozzáállása.

Ezen tényezők közül tökéletesen egyik sem ismerhető meg, és a megismerhetőség foka is különböző. Nehéz, de legalábbis mindenképpen sok munkát igényel megbízható információt szerezni az emberek erkölcséről, és nem feltétlenül szerzünk tudomást arról az ellenőrzés során, hogy valakinek a gyermekét csak többtízmilliós műtéttel lehetne meggyógyítani, valahol a világ másik felén (ami – érthető okokból – elég erős nyomás lehet a sikkasztásra). Ez a kockázat (azaz a vezetők által elkövetett gazdasági bűncselekmény) tehát olyan, melynek bekövetkezési valószínűségét legfeljebb csak nagyságrendileg lehet meghatározni.

A leírtak miatt a kockázat első, valószínűségi dimenziója is legtöbbször csak elnagyolt módon, például alacsony-közepes-magas, vagy hasonló sorrendi skálán határozható meg.

Ha rendelkeznék is pontos mérőszámokkal a kockázatok hatásáról és valószínűségéről, akkor is gondot jelentene a kockázatok rendezése a kétdimenziós térben. Vajon a 0,001% valószínűségű, többmilliárd forintos kockázat, vagy a 10% valószínűségű, pár tízezer forintos kockázat nagyobb? Tovább nehezíti a problémát, hogy általában nem is rendelkezünk pontos mérőszámokkal. Ennek következtében az egyes kockázatok becslése, a kockázati sorrend felállítása nem egzakt tudomány, rengeteg szubjektív elemet tartalmaz, és alapvetően a döntésemélet eszköztárával kezelhető. Noha a rendelkezésre álló döntéseméleti eszköztár

elég bőséges, van pár általánosan elterjedt, „közkezdvelt” elemzési módszer a kockázati sorrend kialakítására, melyet az alábbiakban ismertetek.

Az ellenőrzési szakmában hagyományosnak tekinthető kockázatelemzési eljárás

Az ellenőrzési folyamat (elvileg) első lépése az ellenőrzés céljának/céljainak, tehát végeredményben az igazolandó állításnak a meghatározása. Ezzel összefüggésben kell definiálni, hogy az ellenőrzött állítás mikor tekinthető igaznak, és mikor nem: pontosan ezt jelenti a lényegesség. Az elvárt bizonyosság meghatározása szintén ekkor történik, és csakúgy mint a lényegesség esetén, itt is általában a megbízóval konzultálva³⁸. Fontos kiemelni, hogy a bizonyítékok rendelkezésre állásától függően mind lényegesség, mind az elvárt bizonyosság módosulhat³⁹ a későbbiekben (például azért, mert az eredetileg meghatározott értékek mellett túlságosan költséges volna az ellenőrzést végrehajtani).

Az ISA 315 szabványból kiindulva, azt általánosítva a kockázatelemzési folyamatnak az alábbi főbb szakaszai különböztethetők meg.

- Az ellenőrzött szervezetnek és környezetének tanulmányozása.
- Az ellenőrzés szempontjából különbözőnek tekinthető részterületek azonosítása, pontosabban a létezésükkel, elkülönültségükkel kapcsolatos hipotézisek felállítása, például „kockázati faktorok” definiálásával. Ez gyakorlatilag annyit jelent, hogy elkülönítjük a részterületeket az ellenőrzött állításhoz való kapcsolódásuk módja szerint. Mindez azért lényeges, mert ezen a módon homogén „részsokaságokat” képzünk, ami növeli a későbbi elemzések hatékonyságát.
- Az egyes területek kockázatainak azonosítása, valószínűségük és hatásuk felmérése (ez a szorosabb értelemben vett kockázatelemzés).
- A részterületek kockázatainak azonosítása, és esetleges összegzése után meg kell határozni, hogy a felderített kockázatok fényében melyik területen, milyen további vizsgálatok szükségesek. Ez jelenti mind a vizsgálat módszerét (pl. elemző, tesztelő, azon belül milyen mintavétel), mind a vizsgálat mélységét (pl. mintanagyság stb.).

³⁸ A konzultáció nem feltétlenül jelent közvetlen kommunikációt: az Állami Számvevőszék legtöbb ellenőrzése esetében ez a konzultáció a jogszabályok értelmezését, vagy a közérdeklődés felmérését jelenti.

³⁹ Vannak azonban ellenőrzések, ahol ezek az értékek jogszabályban rögzítettek, így nincs lehetőség módosításra.

- A kockázatelemzés egy iteratív folyamat, ha például a részletek tesztelése, vagy más későbbi munkafázis során felmerülnek új szempontok, akkor azok fényében újra kell értékelni a kockázatokat.

A kockázatelemzés egyik legfőbb módja az elemző eljárások alkalmazása. Ennek oka, hogy az elemző eljárások nagyon hasznosak a „szokatlan” tranzakciók, események, trendek stb. felderítésében. Amikor tehát a kockázatelemzésre elemző eljárásokat használunk, akkor olyan ésszerű feltételezéseket kell gyűjteni, melyek igazak akkor, ha „minden rendben”; ezen feltételezéseket összehasonlítva a tényekkel azután azonosítani lehet a kockázatos pontokat. Nyilvánvaló, hogy minél „durvább”, minél kevésbé részletezett adatok állnak rendelkezésre a kockázatelemzéshez, annál kevésbé lesz pontos az elemzés eredménye is.

Az auditori szakma természetes velejárója, sőt, a szabványokban is lefektetett követelmény az „egészséges” szkepticizmus⁴⁰. Ez a kockázatelemzés során azt jelenti, hogy nem szabad elsiklani semmilyen gyanús dolog felett. Különösen igaz ez a könyvvizsgálatra, ahol a könyvvizsgálónak kötelessége a csalások, bűncselekmények lehetőségét is számba venni.

A kockázatelemzés a legtöbb ellenőrzési alkalmazásban két fő területre összpontosít: az eredendő- és a kontrollkockázatok felmérésére. Fontos azonban megjegyezni, hogy ez az elemzés nem a teljes szervezeti működést érinti, hanem csak azokat a kockázatokat, amelyek az ellenőrzött állítással kapcsolatban vannak. Amikor például egy önkormányzatnál a normatív támogatások felhasználását ellenőrzi az Állami Számvevőszék, akkor a vizsgálat semmiképpen nem tér ki arra, hogy a munkatársak rendszeres késéseit milyen módon, milyen kontrollokkal próbálják megakadályozni a helyi önkormányzatnál.

Az eredendő kockázatok felmérésekor a legfontosabb a szervezetnek és környezetének alapos ismerete/megismerése. Magukat a kockázatokat mindig úgy kell értékelni, mintha nem léteznének kontrollok. Klasszikus eredendő kockázatok egy könyvvizsgálatnál:

- a szervezet típusa, a tevékenység jellege
- a szervezeti működés (in)stabilitása, mióta végzi a szervezet a tevékenységet
- külső környezeti tényezők (piac, kormány), a külső környezet változásai
- a vezetők és alkalmazottak tapasztalata, hozzáértése, képzettsége, a fluktuáció mértéke
- a szervezetben uralkodó etikai normák

⁴⁰ professional skepticism

Az eredendő kockázatok felmérésére legtöbbször űrlap készül, ami egyrészt tartalmazza a kockázatok megnevezését, a bekövetkezés valószínűségét és hatását. Ezek alapján valamilyen módon egy értéket, „kockázati súlyt” rendelnek minden kockázathoz, és ezt valamilyen – általában egy tudományos köntösbe bújtatott ad hoc – módszerrel összegzik az adott területre. Sokszor az így kapott összesített értéket visszafordítják százalékos formára, hogy azt behelyettesíthessék az (1.4) képletbe. Természetesen ez utóbbi eljárás elég gyenge alapokon nyugszik, amint azt az I. részben ki is fejtettem.

A kontrollkockázat felmérése nem magától értetődő dolog, és a gyakorlatban sokszor hibásan is végzik. A kontrollt ugyanis egy részről meg kell tervezni, más részről végre kell hajtani. A kockázat pedig mind a tervezésnél, mind a végrehajtásnál lehet.

A tervezés jósága azt jelenti, hogy az adott kontroll, amennyiben végrehajtódik, be tudja tölteni tervezett funkcióját, és összességében a kontrollok rendszere képes ésszerű módon csökkenteni azon eredendő kockázatokat, amelyek a téves állítás előidézését jelenthetik.⁴¹ Ennek megítélése alapos felkészültséget, sok tapasztalatot és kreativitást igényel. Előnyös a „fordított gondolkodás”, amikor az ellenőr azon kezd el gondolkodni, hogyan tudná (ha egyáltalán) kijátszani a rendszert, hogy az így téves állítást produkáljon. A lehetséges kijátszási módok életszerűségének/életszerűtlenségének vizsgálata által megítélhető, hogy a felállított kontrollok (elméletben) mennyire megbízhatóak.

A végrehajtás jóságának megítélése, pontosabban a megítélésre alkalmazott technika különbözik az automatikus és a manuális kontrollok esetén. Míg az automatikus kontrollok működéséről (elvileg) meg tudunk győződni egyetlen tranzakció vizsgálatával, addig a manuális kontrollok megbízható működését kontrollteszteléssel (részletesen lásd a III. részben) tudjuk megítélni. Az automatikus kontrollról azért csak elvileg tudunk meggyőződni, mert a gyakorlatban nehéz eldönteni, mi az automatikus kontroll. Ahogy terjed az elektronikus adatfeldolgozás, úgy kerül egyre több folyamat manuális rendszerekből automatikus rendszerekbe. Az átmenet azonban csak a legutóbbi esetben tökéletes. A legtöbb informatikai rendszer nem zárt, hozzáértő rossz nyelvek szerint például a hazai piacon nem

⁴¹ Érdemes odafigyelni a definícióra. Ugyanazon szervezetnél teljesen más lehet a kontrollkockázat, ha más az ellenőrzési cél. Ha a cél a beszámoló megbízhatóságának megítélése, a releváns kontrollok egyike lehet például a rendszeres pénztárellenőrzés: a terület rossz szabályozása növeli annak kockázatát, hogy a beszámolóban lényeges hiba van. Abban az esetben, ha az ellenőrzés célja a „pénztárellenőrzés kontroll” megbízhatóságának megítélése (egy tipikus belső ellenőri feladat), akkor ez esetben a rosszul szabályozott kontroll nem jelent kontrollkockázatot, hiszen pont a működés megfelelőségének megítélése az ellenőrzés célja, ez az igazolandó állítás („a pénztárellenőrzés jól működik”).

kapható olyan számviteli szoftver, amellyel ne lehetne (megfelelő jogosultság birtokában) utólag beleírni APEH által lezárt időszak könyvelésébe. Sokszor felkészült informatikai auditorokra van szükség annak elbírálásához, hogy az „automatikus” kontrollok mennyire megbízhatóak.

Állandó vita az ellenőri szakmában, közvetlen környezetemben is, hogy vajon érdemes-e olyankor tesztelni a kontrollok működését, ha azok nincsenek megfelelően szabályozva. Az egyik álláspont szerint nem: ha ugyanis nincs szabályozás, valójában a kontrollok, azok egymásra épülésének „tervrajza” sincs dokumentálva, így nincs mihez viszonyítsunk a teszteléskor. Ez az óvatosabb álláspont, ugyanis ilyenkor maximális kontrollkockázatot feltételezve fogunk további eljárásokat végezni. A másik álláspont szerint a kontrollok rendszerét nem feltétlenül szabályzatokkal lehet kialakítani, közvetítheti azt a szervezeti kultúra, esetleg szóbeli utasítás is: a lényeg, hogy működjön a kontroll. Véleményem szerint az első álláspont követése a célszerűbb, a második esetben ugyanis teljesen kiszámíthatatlan kockázatoknak teszi ki magát az ellenőr.

A kontrollkockázat felmérése, hasonlóan az eredendő kockázathoz, általában űrlapon történő pontozással, összesúlyozással történik. Nem sikerült olyan módszertani munkát olvasnom, ahol explicit javaslatokat tettek arra, hogyan kell figyelembe venni az egyes kockázatok közötti korrelációt. Ennek megfelelően a kontrollkockázatra kapott valószínűségi értéket ismételten behelyettesítve az (1.4) képletbe, megkapjuk, hogy az adott ellenőrzési bizonyosság mellett mekkora feltárási valószínűség kell az egyes részterületeken, és ezáltal mekkora mintát szükséges venni az ellenőrzés későbbi szakaszában. Az egyetlen meglepő dolog ebben, hogy ez a módszer – minden hiányossága ellenére – a tapasztalatok szerint mégis működik, legalábbis, eddig még nem sokan panaszkodtak. De ez egy másik kérdés, aminek tárgyalása nagyon messzire vezet.

Az ellenőrzésben használt kockázatelemzési módszerek továbbfejlesztési lehetőségei

A most bemutatott „hagyományos” módszer – bizonyos feltételek fennállása esetén – tovább finomítható: a szakasz hátralévő részében erre teszek pár javaslatot. A javaslatok gyakorlati tesztelésére sajnos nem volt lehetőségem, az egy későbbi időszak feladata lesz. Mindennek ellenére – a későbbi kutatások irányának kijelölése miatt is – indokoltnak tartom a javaslatok ismertetését.

Az eredendő és a kontrollkockázat kapcsolata

Amennyiben az eredendő és a kontrollkockázatok nem függetlenek, az (1.4) képlet közvetlenül nem alkalmazható. Az ebből származó probléma egyik lehetséges feloldása az lehet, ha ahelyett, hogy közvetlenül az eredendő és a kontrollkockázatot határoznánk meg, és „szoroznánk össze”, külön-külön megvizsgáljuk az eredendő kockázat összetevőit, és azok viselkedését vetjük össze külön-külön a kontrollokkal. Amennyiben az egyes eredendő kockázati tényezők „nem kivédett” részei (a 2. ábra utolsó, „végeredmény” „oszlopa”) függetlenek, akkor a komplementer események „szorzása” innentől elvégezhető. Ha a függetlenség nem áll fenn, de a hasonló szervezeteknél szerzett tapasztalatok alapján stabilnak mutatkozik a korrelációs struktúra, ezt a struktúrát felhasználva szintén meghatározható az eredendő és a kontrollkockázat együttes mértéke.

Az eredendő és a kontrollkockázat együttes mértékét az (1.2) képletben leírt módon határozhatjuk meg, amit most a könnyebbség kedvéért megismételek:

$$\Pr\left(\bigcup_i \left(IR_i \cap \bigcup_j CR_{ij} \right)\right). \quad (2.1)$$

Ha megvizsgáljuk a képletet, azonnal látszik a fenti javaslat mögötti intuíció: az

$$IR_i \cap \bigcup_j CR_{ij} \quad (2.2)$$

tagok jelképezik az egyes eredendő kockázati tényezők „nem kivédett” részeit. Ha ezek függetlenek, akkor uniójuk valószínűségét a De Morgan azonosságok alapján egyszerű szorzással megkaphatjuk. Ha a függetlenség nem áll fenn, viszont a korrelációs struktúra ismert, szintén egyszerűen meghatározható az együttes kockázat.

A későbbi kutatások feladata ezen a téren kettős: vagy elégséges bizonyítékot kell szerezni arra vonatkozóan, hogy a Magyarországon a közsférában a kockázat (2.2) képlettel számított tényezői egymástól függetlennek (korrelálatlannak) tekinthetőek, másrészt ha ezt nem sikerül bizonyítani, akkor meg kell győződni arról, hogy a korrelációs struktúra tekinthető-e annyira stabilnak, hogy arra a későbbiekben lehessen alapozni.

Ha a kutatás eredményeképpen az derül ki, hogy sem a függetlenség, sem a korrelációs struktúra stabilitása nem áll fenn, akkor más lehetőség hiányában a legjobb megoldás minden egyes eredendő kockázati tényező (IR_i) bekövetkezési valószínűségét 100%-nak tekinteni.

Ebben az esetben ugyanis egyrészt biztosan felülbecsüljük a kockázatot, másrészt a (2.1) képletből kiesik az eredendő kockázat, ami így a kontrollkockázatok uniójára egyszerűsödik.

A kockázati tényezők aggregálási módja

Akár figyelembe vesszük az eredendő és a kontrollkockázat kapcsolatát, akár nem, mindenképpen szembesülünk a kockázati tényezők aggregálásának problémájával. Jelenleg – hasonlóan a kockázati tényezők értékeléséhez – az aggregálás is teljesen szubjektív módon, egy lineáris modellben történő súlyozás segítségével történik. A súlyok meghatározásának módja jobb esetben valamilyen kollektív értékelési technika, amikor is több szakember összeül, és konszenzussal kialakítja az egyes tényezők súlyát. Nem szükségszerű azonban sem a lineáris összegzés, sem a súlyok ad hoc meghatározása.

Amennyiben ugyanazt a területet folyamatosan vizsgáljuk (tehát sok megfigyeléssel rendelkezünk), továbbá mind a kockázati faktorok, mind a kockázat hatása könnyen mérhető (vagy mérhetővé tehető), akkor lehetőség van arra, hogy valamilyen jól megválasztott regressziós modellel becsüljük a kockázati súlyokat, illetve különböző modellek illeszkedését összehasonlítva finomítsuk a használt modellünket. Jelenleg ilyen területnek tekinthető az EU támogatások ellenőrzése, ahol az intézményrendszer relatív állandósága és az ellenőrzések nagy száma miatt elegendő megfigyelés áll rendelkezésre ahhoz, hogy egy – a jelenleginél megbízhatóbb – kockázatelemzési modellt lehessen felállítani és kalibrálni⁴². Sajnálatos módon a szükséges adatokhoz való hozzáférés korlátozott, így ezt a feladatot nem állt módomban elvégezni.

A számjegyelemzés, mint speciális kockázatelemzési módszer

A 240-es ISA standard részletesen foglalkozik azzal, hogy a könyvvizsgálat során nem csak a véletlen, vagy hanyagságból fakadó hibákat kell figyelembe venni, hanem oda kell figyelni a potenciálisan előforduló „szándékos” hibákra, tehát csalásokra és egyéb gazdasági bűncselekményekre is. Ez általánosítva is igaz, nem csak a könyvvizsgálat esetén: az auditálandó állítás hamissága nemcsak vétlen tévedésből fakadhat, hanem szándékos manipulációból is. Amikor tehát a lényeges hiba kockázatát felmérjük, figyelembe kell venni

⁴² Egy, az EU támogatásokhoz kapcsolódó lineáris kockázatelemzési modell leírását lásd Lolbert [2003].

azt is, hogy részterületenként eltérő lehet a manipulációból származó eltérések kockázata. Ez a rész egy olyan módszert ismertet, amely segíthet ezt a kockázatot felmérni.

Az egyik legújabb elemző eljárást számjegyelemzésnek (angolul: digital analysis) hívják, alkalmazása csak körülbelül egy évtizede kezd teret nyerni az ellenőrzésben. Maga az eljárás egy megfigyelésen alapul: a nem mesterségesen generált számhalmazokban az egyes pozíciókban (pl. első, második, első kettő, utolsó, utolsó kettő stb.) lévő számjegyek eloszlása nem lehet akármilyen, hanem egy Benford-törvénynek nevezett szabályt követ. Konkrétan az első számjegyek eloszlása tízes számrendszerben a következőképpen adható meg a Benford-törvény szerint:

$$\Pr(x \text{ első számjegye} = d) = \log_{10} \left(1 + \frac{1}{d} \right), \text{ ahol } d = 1, 2 \dots 9 \quad (2.3)$$

Ugyanez számszerűsítve:

Első számjegy	Relatív gyakoriság
1	30,10 %
2	17,61 %
3	12,49 %
4	9,69 %
5	7,92 %
6	6,69 %
7	5,80 %
8	5,12 %
9	4,58 %

3. Táblázat: az első számjegyek relatív gyakorisága tízes számrendszerben

Ez a jelenség azért használható fel az ellenőrzésben, mivel a manipulált adatokat tartalmazó számhalmazok már nem az előírt eloszlást követik. A következőkben először röviden ismertetem a Benford-törvény történetét, továbbá az irodalomban fellelhető intuitív, statisztikai és valószínűségelméleti magyarázatokat. Ezt követően bemutatok egy – meglehetősen matematika-intenzív – bizonyítást a Benford-törvényhez, amiből kiindulva egy, a gyakorlat szempontjából is releváns magyarázatot adok arra, hogy empirikusan miért tapasztaljuk mindenhol a törvény fennállását. Mind a bizonyítás, mind az ezt követő magyarázat azon alapul, hogy a Benford törvény kielégítése egy adott számrendszerben egyenértékű azzal, ha a valószínűségi változó eloszlás egy transzformáltjának karakterisztikus függvénye adott (a számrendszer alapjától függő) pontokban nulla értéket vesz fel. Végül pedig egy részletes áttekintést adok a gyakorlatban használt tesztekéről, amit felhasználva bármely, statisztikához kicsit értő ellenőr már végrehajthatja a szükséges elemzéseket.

A Benford-törvény története

1881-ben Simon Newcomb egy American Journal of Mathematics-ben megjelent cikkében publikálta azt a régi megfigyelést, hogy a könyvtárban található logaritmus táblázatok első oldalai sokkal elhasználtabbak az utolsó oldalaknál. Noha ez a jelenség természetes volna egy rossz regény esetén, semmiképpen sem gondoljuk természetesnek egy matematikai, fizikai, mérnöki számításokhoz használt eszköznél. Newcomb a megfigyelést alapul véve megadja a (2.3) képletet.

Ennek a publikációnak nem volt nagy hatása, olyannyira, hogy 1938-ban Frank Benford újra publikálhatta ugyanezt a jelenséget, mint új eredményt. Benford ugyanazzal a logaritmustáblázatos megfigyeléssel kezdi cikkét, mint Newcomb, azonban ő továbbmegy ennél: az élet különféle területeiről vett adathalmazokon teszteli megfigyelését, például 335 folyó felszínén, 3259 település lakosságszámán, a kémiai elemek mol-tömegein, a természetes számok hatványain, egy hírességekről szóló magazinban felsorolt híres emberek házszámain, és így tovább, mindösszesen 20229 adaton. Megfigyelései alapján ugyanúgy levezeti a (2.3) képletet, mint Newcomb, azonban ennél többet is állít: a törvényszerűség igaz marad akkor is, ha az adatok reciprokát vesszük, és akkor is, ha a számokat más számrendszerben írjuk fel (ekkor a logaritmust 10-es alap helyett a számrendszer alapijával kell venni). Mivel az ő cikke egy „fontos” cikk mellett jelent meg, ezért Newcomb cikkével ellentétben nem került el a tudós társadalom figyelmét. A nevezett jelenséget tehát 1938 óta általában Benford-törvénynek hívják⁴³.

Lehetséges magyarázatok a törvényszerűsége

Érdemes leszögezni, hogy a Benford-törvény semmiképpen sem nevezhető univerzális törvényszerűségnek. Tekintsük például a felnőtt emberek méterben mért testmagasságát: nyilvánvaló, hogy az adatok kb. 99%-a 1-es számjeggyel kezdődik, hiszen kevés a 2 méternél magasabb, illetve 1 méternél alacsonyabb felnőtt. Vagy: egy adott város telefonszámai általában nem kezdődhetnek akármilyen számjeggyel. Mindezen nyilvánvaló ellenpéldák dacára több kezdeti szerző szerint maga a számok jelenleg elterjedt írási módja, a pozicionális

⁴³ Elterjedt nevek még: Newcomb-Benford-törvény, „első számjegy törvény”, „szignifikáns számjegy törvény”.

írási mód (szemben például a római számokkal) okozná a jelenséget. Később természetesen születtek ennél kifinomultabb magyarázatok is.

Vannak olyan egyszerű példák, empirikus megfigyelések a törvény érvényesülésére, amelyek nehezen tudhatóak be egyszerű véletlennek. Az első ezek között a szorzótábla, ahol az „1” számjegy jóval gyakrabban kezdő számjegy, mint ahogyan azt várnánk:

	1	2	3	4	5	6	7	8	9
1	1	2	3	4	5	6	7	8	9
2	2	4	6	8	10	12	14	16	18
3	3	6	9	12	15	18	21	24	27
4	4	8	12	16	20	24	28	32	36
5	5	10	15	20	25	30	35	40	45
6	6	12	18	24	30	36	42	48	54
7	7	14	21	28	35	42	49	56	63
8	8	16	24	32	40	48	56	64	72
9	9	18	27	36	45	54	63	72	81

4. Táblázat: az „1” számjegy megjelenése kezdő számjegyként a szorzótáblában

Látható, hogy az „1” ugyan nem a Benford-törvény által előre jelzett (30,1%) arányban fordul elő, de ennek ellenére sokkal gyakrabban, mint bármely más kezdő számjegy, és konkrétan kétszer olyan gyakran, mint azt egyenletes eloszlás esetén várnánk. Figyeljük meg, hogy a „peremeloszlások” a szorzótábla szélén még egyenletesebbek! Egyébként megmutatható, hogy a szorzótábla általánosított verziójában, tehát amikor nem kettő, hanem n db, az $[1; 10)$ intervallumba eső (tehát nem csak egész) számot szorzunk össze, akkor a kezdőszámjegyek relatív gyakorisága tart a Benford-törvény által előre jelzett értékekhez.

A leginkább szemléletes gyakorlati példa éppen a közgazdaságtanból származik. Azt fogjuk ugyanis megmutatni, hogy egy jól körülírható, és nem is túlságosan restriktív feltevés esetén a bankszektorban fennálló összes egyedi számlaegyenleg kezdő számjegyei a Benford-törvénynek megfelelően oszlanak el tetszőleges pillanatban, legalábbis, ha nincs pénzkivétel.

Tegyük fel, hogy 1000 forintot teszünk a számlánkra 10% kamatláb mellett. Több, mint 7 évbe telik, amíg az összeg 2000 forintra „hízik”. A 3000 forint eléréséhez már csak kb. 4 év kell, és könnyű belátni, hogy a legrövidebb idő a 9000 és 10 000 forint között telik el. A 10 ezer és 20 ezer forint között viszont ismét 7 év telik el. Tehát amennyiben a konkrét számla egyenlegét nézzük az időben, azt fogjuk tapasztalni, hogy az idő legnagyobb részében az 1-

essel kezdődik, és 9-cel fog a legrövidebb ideig kezdődni. Nyilvánvalóan ez a kijelentés igaz akkor is, ha nem 1000 forintot teszünk be, vagy nem 10% kamatot használunk. Ha olyan feltevéssel élünk, hogy a bankszámlaképződés folyamata ergodikus⁴⁴, akkor bármely időpontban igaz ez a „keresztmetszeti” adatokra. Ez a logikát felhasználva általánosan is kijelenthető, hogy minden $a \cdot q^n$ geometriai sorozatra (q nem 10 hatvány), illetve minden $f(x) = a \cdot q^x$ függvényre igaz a Benford-törvény.

A közgazdaságtannal és a társadalmi jelenségekkel összefüggő adatok jelentős része lognormális, exponenciális, vagy ezekhez hasonló ferde eloszlást követ. Intuitíve is jól látható, hogy ilyen eloszlásból származó adatoknál magasabb lesz a kisebb kezdő számjegyek gyakorisága. A pontos valószínűség kiszámítása ugyan elég körülményes is lehet, a szimulációval kapott értékek nagyon közel állnak a Benford-törvény által adott relatív gyakoriságokhoz.

A Benford-törvényt szokás úgy interpretálni, hogy „annak valószínűsége, hogy egy véletlenszerűen kiválasztott szám d -vel kezdődik, $\log_{10}(1 + 1/d)$ ”. Noha ismeretterjesztő jellegű munkákban ez a megfogalmazás tökéletesen megfelelő, valójában rossz definíció. Sem nem ismerünk ugyanis módszert arra, hogy véletlenszerűen kiválasszunk egy számot, sem pedig nem tudunk a kezdő számjegyekre vonatkozóan ahhoz hasonlóan megadni természetes relatív gyakoriságot az összes egész szám felett, mint például meg tudunk adni a páros számokra (0,5) vagy például az n -nel osztható számokra ($1/n$) nézve.

Elméletileg az ilyen valószínűségeket úgy kellene kiszámítani, hogy 1 és n közötti számokra meghatározzuk a relatív gyakoriságot, majd vesszük ennek határértékét, ahol n tart a végtelenbe. A probléma az, hogy ez az elv nem alkalmazható, ugyanis a kérdéses határérték – a klasszikus értelmezés szerint legalábbis – nem létezik. A klasszikus definíció szerint nem konvergens sorozatok határértéke megadható különféle összegző módszerekkel. Az összegző módszerek kiterjesztik a határérték fogalmát, azonban ez a kiterjesztés nem feltétlenül egyértelmű. A Flehinger [1966] által definiált összegző módszer például a Benford-törvénynek megfelelő értékeket ad vissza, míg a Raimi nagy összefoglaló művében (Raimi [1976]) Stigler-törvény néven idézett módszer más értékeket ad.

Mivel az összegző módszerek csak ad hoc magyarázatot kínáltak a jelenségre, a magyarázatok egy új családja is megjelent, nevezetesen az invariancia-elven alapuló magyarázat. Ez

⁴⁴ Az idősorok ergodikus voltát például az idősorelemzésben rendszerint fel szokás tételezni.

valójában nem is magyarázza a jelenséget, sokkal inkább körülhatárolja azon eseteket, amikor az adatok számjegyeinek eloszlása a Benford-törvényben előírtak megfelel.

A legegyszerűbb invariancia a skála-invariancia, amelyet (Pinkham [1961]) definiál. Az ötlet onnét származik, hogy az empirikus adatok továbbra is eleget tesznek a Benford-törvénynek, ha pl. méterből lábba, vagy yardba váltják át őket, vagy az értékbeni adatokat dollárból euróba. A skála invarianciához hasonló ihletésű az eltolás-invariancia (Raimi [1969]), ez például a hőmérsékleti skálák közötti műveleteknél lehet érdekes. Végül a legáltalánosabb invariancia a bázis-invariancia, amely azt mondja ki, hogy bárhogyan is írjuk fel más számrendszerben az adatokat, például az első számjegy valószínűségére igaz lesz a

$$\Pr(x \text{ első számjegye} = d) = \log_b(1 + 1/d), \text{ ahol } d = 1, 2, \dots, b-1. \quad (2.4)$$

Kicsit más jellegű tulajdonság az összeg-invariancia⁴⁵: ez azt jelenti, hogy egy Benford-törvénynek engedelmeskedő számhalmazban az adott számjegy-sorozattal kezdődő számok összegének várható értéke csak attól függ, hogy milyen hosszú a számjegysorozat. Konkrétan, a 9-el kezdődő számok összege megegyezik a 4-el vagy az 1-el kezdődő számok összegével, azaz várható értékben a főösszeg 1/9 része. Hasonlóan, a 352-vel és 815-tel kezdődő számok összege is megegyezik, ezek a részösszegek a főösszeg 1/900 részét adják várható értékben.

Az invariancia-elvekkel kapcsolatos eddigi eredmények, illetve saját eredményeim a (Lolbert [2006b]) cikkben olvashatóak. Mivel azonban egyik invariancia-elv sem szolgál valós magyarázatként a Benford-törvényre, ezért ismertetésük nem szolgálná a disszertáció céljait.

Az irodalomban található magyarázatok leginkább működőképes csoportja azt a jelenséget használja, amely már Benford eredeti cikkében is megfigyelhető volt: az egyes adatsorok ugyan külön-külön csak nagyjából követték az előírt eloszlást, az adatok összességére nézve viszont nagyon jó volt az illeszkedés. Ennek a jelenségnek egy külön fejezetet szentel a Benford-törvényről szóló első alapvető írás⁴⁶ is, felsorolva több jelentős eredményt.

Az eloszlások keverésével összefüggő magyarázatot korrekt formában Hill [1996] publikálta. Az általa kimondott tétel szerint, ha véletlenszerűen kiválasztott eloszlásból veszünk véletlen számot, akkor az így kapott szám számjegyeinek eloszlása a Benford-törvénynek megfelelő lesz. A valós élet adatsorai tényleg ilyenek. A cégek nyeresége például sok tényezőtől függ: kezdve a gyártott termékek/szolgáltatások iránti kereslettől, egészen a felhasznált erőforrások

⁴⁵ Allaart, [1997]

⁴⁶ Raimi, [1976]

aráig. Ebben az esetben rengeteg véletlen számot látunk, amelyeknek valamilyen kombinációja adja majd a nyereséget.

A következő oldalakon kísérletet teszünk egy új magyarázattal szolgálni a jelenségre.

A Benford-törvény egy új magyarázata

A most következő formalizmus nem öncélú: ezek a tételek teszik majd indokolttá a későbbi alkalmazásokat, legalábbis azok egy részét.

Az első pár definíció alapvetően technikai jellegű, a későbbi tárgyalást könnyíti meg.

1. definíció (mantissza függvény)

Minden $b > 1$, $b \in \mathbb{N}$ esetén b alapú mantissza függvénynek nevezzük az alábbi hozzárendeléssel megadott $M_b : \mathbb{R}^+ \rightarrow [1, b)_{10}$ függvényt: $M_b(x) = x \cdot b^{-[\log_b x]}$. Az $M_b(x)$ értéket x b alapú mantisszájának hívjuk. Alapértelmezésben a mantissza alaki értéke b alapú számrendszerben értelmezett.

Példa

A mantissza függvény valójában csak a „tizedesvesszőt” tolja el úgy, hogy a kapott érték 1 és b közé essen. $M_{10}(e) = M_{10}(100 \cdot e) = M_{10}(e/100) = e$.

2. definíció (k-hosszúságú szignifikáns-számjegy sorozat)

Az i -edik k -hosszúságú szignifikáns-számjegy sorozat alatt értjük azt az $S_{i,k}^{(b)}(x)$ függvényt, amely az x valós számhoz hozzárendeli az $M_b(x)$ mantissza b^{-i+1} pozíciótól b^{-i-k+2} pozícióig terjedő k darab számjegyét. Alapértelmezésben $k = 1$, ilyenkor elhagyjuk az alsó indexet, csakúgy, mint a (b) felső indexet, ha nem akarjuk hangsúlyozni azt a számrendszert, amit használunk.

Megjegyzés az 2. definícióhoz

A mantissza definíciója miatt $S_1^{(b)}(x) \in [1, b)$, tehát az első számjegy nem lehet nulla.

Példa

$$S_1(\pi) = S_1(\pi/100000) = 3, S_2(1999.\dot{9}) = S_2(0.02) = 0, S_{1,3}(\pi) = 314, S_1^{(3)}(\pi) = 1.$$

A „vegyünk egy véletlen számot” kitételt sem Newcomb, sem Benford nem definiálta precízen. Mivel minket a szignifikáns számjegyek eloszlása érdekel, ezért most definiáljuk azt a minimális szigma-algebrát, ami szerint $S_{i,k}^{(b)}(x)$ mérhető.

3. definíció (b alapú mantissza-sigma-algebra)

Tetszőleges $E \in \mathcal{B}([1, b))$ halmazra legyen $\mathcal{M}_b(E) = \bigcup_{n \in \mathbb{Z}} b^n E$. A következőben b -alapú mantissza-sigma-algebrán az $\mathcal{M}_b = \{\mathcal{M}_b(E) : E \in \mathcal{B}([1, b))\}$ halmazrendszert értjük, ahol $\mathcal{B}(A)$ jelöli az A -n értelmezett Borel halmazokat.

Megjegyzés a 3. definícióhoz

Könnyen látható, hogy $\mathcal{M}_b$ szigma-algebra, éspedig pontosan az, amit az $M_b(x)$ mantissza-függvény generál. Az is könnyen látható, hogy $\mathcal{M}_b$ -nek nem eleme a korlátos halmazok, ezért az „összes pozitív valós szám mediánja”-val kapcsolatos probléma, melyet többek között Raimi [1969] is idéz, nem léphet fel. Valójában $\mathcal{M}_b$ összes nemüres részhalmaza log-periodikus, emiatt pedig korlátlan, és torlódási pontja van a 0-ban és a plusz végtelenben.

Példa

$$\mathcal{M}_b(\{1\}) = \{b^n : n \in \mathbb{Z}\}, \mathcal{M}_b([1, b)) = \mathbb{R}^+$$

4. Lemma ($\mathcal{M}_b$ tulajdonságai)

Tetszőleges b bázis mellett az alábbiak igazak $\mathcal{M}_b$ -re:

- a) $\mathcal{M}_b$ zárt a skalárral való szorzásra, tehát tetszőleges $A \in \mathcal{M}_b$ és $\alpha > 0$ esetén $\alpha A \in \mathcal{M}_b$.
- b) $\mathcal{M}_b$ zárt az egész gyökvonásra, tehát tetszőleges $A \in \mathcal{M}_b$ és $a \in \mathbb{N}^+$ esetén $A^{1/a} \in \mathcal{M}_b$.
- c) $\mathcal{M}_b(E) = \mathcal{M}_{b^n}\left(\bigcup_{k=0}^{n-1} b^k E\right) = \bigcup_{k=0}^{n-1} \left(\mathcal{M}_{b^n}(b^k E)\right)$ minden $n \in \mathbb{N}$ és $E \in \mathcal{B}([1, b])$ esetén.
- d) $\mathcal{M}_b \subset \mathcal{M}_{b^n} \subset \mathcal{B}(\mathbb{R}^+)$ minden $n \in \mathbb{N}$ -re.

Megjegyzés a 4. lemmához

A lemma d) pontjában szereplő tartalmazás szigorú. Ennek belátásához tekintsük például az $\mathcal{M}_{100}([2, 20])$ halmazt, ami biztosan nincs benne $\mathcal{M}_{10}$ -ben. Vegyük észre, hogy szemben a skalárral való szorzással, a skalárral való eltolásra nem zárt $\mathcal{M}_b$. Tekintsük például az $\mathcal{M}_{10}(\{1\})$ halmaz eltolását 1-el: az $\mathcal{M}_{10}(\{1\}) + 1 = \{\dots; 1,01; 1,1; 2; 11; 101; \dots\}$ halmaz biztosan nem eleme $\mathcal{M}_{10}$ -nek. (Ne feledjük, hogy egy mantissza-sigma-algebra tetszőleges elemének torlódási pontja van a nullánál!)

Bizonyítás

Az állítások közvetlenül levezethetők a definíciókból. □

5. definíció (mantissza-eloszlás)

Tetszőleges, $\mathcal{M}_b$ -t tartalmazó szigma algebrán értelmezett $\mathcal{P}$ valószínűségi mértékhez és b bázishoz rendeljük hozzá a $([1, b], \mathcal{B}([1, b]))$ mérhető téren definiált $\mathcal{P}_b$ valószínűségi mértéket úgy, hogy tetszőleges $E \in \mathcal{B}([1, b])$ halmazra $\mathcal{P}_b(E) = \mathcal{P}(\mathcal{M}_b(E))$. Ezt a $\mathcal{P}$ mantissza-eloszlásának hívjuk a b bázisú számrendszerben.

6. definíció (Benford-törvény)

Azt mondjuk, hogy egy $\mathcal{P}$ valószínűségi mérték a b bázisú számrendszerben kielégíti a Benford-törvényt, ha tetszőleges $x \in [1, b)$ esetén $\mathcal{P}_b([1, x)) = \log_b x$.

Megjegyzés a 6. definícióhoz

Benford törvénye kettes számrendszerben ($b=2$) azt jelenti, hogy az 1-essel kezdődő mantissza valószínűsége $\log_2 2 = 1$, ami nem meglepő, figyelembe véve, hogy a kettes számrendszerben csak 1-essel kezdődhet szám. Amennyiben adott a számrendszer bázisa, könnyen levezethetőek speciális törvények az n . számjegyek, vagy számjegycsoportok eloszlására is, ahogy erre példákat is fogunk látni. Érdeemes felhívni a figyelmet arra, hogy az m . és n . pozícióban lévő számjegyek eloszlása nem független, a függőség azonban csökken, minél „hátrébb” megyünk.

Példa

A Benford-törvény szerint tízes számrendszerben annak a valószínűsége, hogy egy tetszőleges szám 1 0 2 számjegyekkel kezdődik, pontosan megegyezik az $[1,02, 1,03)$ intervallum mértékével, ami $\log_{10} 1,03 - \log_{10} 1,02 = \log_{10} (1 + 1/102) \approx 0,42\%$.

Annak valószínűsége, hogy a második számjegy d , kiszámolható az 1 d , 2 d , ... 9 d számjegycsoportokkal való kezdődés összegeként: $\sum_{i=1}^9 \log_{10} (1 + 1/(10i + d))$.

A következő tétel azt mondja ki, hogy nem létezhet olyan valószínűségeloszlás, amely minden számrendszerben kielégíti a Benford-törvényt: mivel azonban a gyakorlatban a 16-os alapnál nagyobbakat nemigen szokás használni, könnyen feloldható a tétel és az empiria közötti látszólagos ellentét. Elemi bizonyítás található a tételre Knuth [1981] könyvében, de az ott adott bizonyítás nem világít rá a jelenség azon oldalára, amit ki szeretnénk használni. Az itt ismertetendő bizonyítás a karakterisztikus függvény viselkedését veszi alapul egyes, kiemelt pontokban, és ezzel lehetővé teszi, hogy a Lévy-féle folytonossági tételt alkalmazva újabb magyarázatot adjunk a Benford törvényre.

7. tétel

Nem létezik olyan pozitív tartójú $\mathcal{P}$ valószínűségi eloszlás a pozitív félegyenes Borel-halmazain, melyre minden $b > 1$ és tetszőleges $x \in [1, b)$ esetén $\mathcal{P}_b([1, x)) = \log_b x$.

Bizonyítás

Tegyük fel, hogy létezik ilyen eloszlás, és legyen $\xi > 0$ egy ilyen eloszlású valószínűségi változó. Ilyenkor nyilván létezik az $\ln \xi$ transzformált valószínűségi változó is. Könnyen látható, hogy ennek a transzformált változónak az eloszlása olyan, hogy tetszőleges $b > 1$, $b \in \mathbb{N}$ és $E \in \mathcal{B}([0, \ln b))$ esetén

$$\Pr\left(\ln \xi \in \bigcup_{k \in \mathbb{Z}} (E + k \cdot \ln b)\right) = \frac{\lambda(E)}{\ln b}, \quad (2.5)$$

tehát a Lebesgue-mértéknek egy módosítása. Ehhez egyszerűen ki kell használni azt, amit az eredeti változó mantissza-eloszlásáról feltettünk. Megmutatjuk, hogy ilyen eloszlású valószínűségi változó nem létezhet, és ezzel a tételt be is láttuk.

Amennyiben létezik egy eloszlás, akkor létezik annak karakterisztikus függvénye is. A (2.5) formula megadja minden periodikus halmaz valószínűségét: ez pontosan a halmaz ismétlődő részének és a periódus hosszának aránya. Ez azt jelenti, hogy a karakterisztikus függvény értékét minden $t = 2\pi/\ln b$, $b > 1$, $b \in \mathbb{N}$ pontban meg tudjuk határozni. Mivel a teljes perióduson nézve mind a szinusz, mind a koszinusz függvény integrálja nulla, ezért ezekben a pontokban a karakterisztikus függvény is nulla lesz.

Látható, hogy a karakterisztikus függvény határértéke az origóban ($b \rightarrow \infty$) 0, ugyanitt a felvett értéke pedig (definíció szerint) 1. Mivel a karakterisztikus függvénynek mindig folytonosnak kell lennie, ezért ilyen eloszlás nem létezik. $\square$

Megjegyzések a bizonyításhoz

Nem nehéz megmutatni, hogy nemcsak a b alapú Benford-törvény fennállásából következik, hogy a változó logaritmusának karakterisztikus függvénye a $t = 2\pi/\ln b$ alakú pontokban nulla, hanem ennek az állításnak a megfordítása is igaz (ennek bizonyítása megtalálható Lolbert [2007] 3.2 pontjában). Ha tehát egy változó logaritmusának karakterisztikus függvénye bizonyos kiemelt pontokban nulla, akkor biztosak lehetünk abban, hogy a megfelelő bázisban kielégíti a Benford-törvényt.

Az a tény, hogy a megfigyelések szerint a valóságban olyan sok eloszlás kielégíti többé-kevésbé a Benford törvényt, két okra vezethető vissza. Az egyik ok a Lévy-féle folytonossági tétel, ami leegyszerűsítve arról szól, hogy „közeli” karakterisztikus függvény egyben „közeli” eloszlásfüggvényt is jelent. A másik ok, hogy a gyakorlatban megtalálható eloszlások logaritmikus transzformáltjainak karakterisztikus függvényei ténylegesen gyorsan tartanak a nullához, és az első szükséges helyen már elég közel vannak hozzá. A kettő együttes fennállása azzal jár, hogy a Benford törvény jó közelítéssel érvényesülni fog.

Abban az esetben, ha b korlátos (például legfeljebb 16), akkor már lehet esély arra, hogy teljesüljön a Benford-törvény nemcsak egy, de minden, a korlát értékét nem meghaladó bázisban. Ismert ugyanis, hogy ha egy változónak létezik sűrűségfüggvénye, akkor a hozzá tartozó karakterisztikus függvény $\pm\infty$ -ben tart a nullához. Ebből egyébként következik az is, hogy tetszőleges t pontban tetszőlegesen közel vihetjük egy ilyen η valószínűségi változó karakterisztikus függvényét nullához csak azzal, hogy egy megfelelően nagy konstanssal beszorozzuk a változót. (Ehhez azt kell felhasználni, hogy $\varphi_{a\eta}(t) = \varphi_{\eta}(at)$, és $\lim_{t \rightarrow \pm\infty} \varphi_{\eta}(t) = 0$.) Ezt felhasználva, ha felülről korlátos azon bázisok halmaza, ahol „elvárjuk” a jelenséget, akkor kijelenthető, hogy *tetszőleges pozitív tartójú (sűrűségfüggvénnyel rendelkező) eloszlást megfelelően magas hatványra emelve (az $a \cdot \ln \xi$ változót „visszatranszformálva” ξ^a -t kapunk) tetszőlegesen közel juthatunk a Benford-törvényhez, minden elvárt bázis esetén.* (A hatványozással kapcsolatosan hasonló megfigyelést tett Adhikari-Sarkar [1968] és Adhikari [1969] is.)

Példák

Tekintsünk példaképp két olyan – közhismert és gyakran használt – eloszlást (lognormális, exponenciális), amelyek a tapasztalatok szerint jól illeszkednek a Benford-törvényhez, és vizsgáljuk meg, hogy a logaritmusuknak milyen a karakterisztikus függvénye.

A lognormális eloszlás logaritmusa normális eloszlású, aminek a karakterisztikus függvénye

$$\varphi(t) = e^{it\mu} e^{(-0,5 \cdot \sigma^2 \cdot t^2)} \quad (2.6)$$

Felső becslést adhatunk ennek abszolút értékére:

$$|\varphi(t)| = \left| e^{it\mu} e^{(-0,5 \cdot \sigma^2 \cdot t^2)} \right| = e^{(-0,5 \cdot \sigma^2 \cdot t^2)} \quad (2.7)$$

Ha $t = 2\pi/\ln 10 \approx 2,7$, akkor (2.7) értéke (persze a szórástól függően) elég közel lesz a nullához (a Benford-törvény szerinti hipotetikus értékhez tízes számrendszer esetén), ami érthetővé teszi a jó illeszkedést.

Az exponenciális eloszlás logaritmikus transzformáltjának karakterisztikus függvénye:

$$\varphi(t) = \int_0^\infty e^{it \ln x} \lambda e^{-\lambda x} dx = \lambda^{-it} \Gamma(it+1) \quad (2.8)$$

Ennek abszolút értéke független a λ paramétertől, hiszen mivel t valós, $|\lambda^{-it}| = 1$. A gamma függvény abszolút értéke viszont 2,5-nél nagyobb abszolút értékű t esetén már 0,1-nél közelebb van a nullához, így nem meglepő, hogy az exponenciális eloszlás ilyen jól illeszkedik a Benford-törvényhez.

Korábban volt szó róla, hogy a felnőtt emberek méterben mért testmagasságának eloszlása „rosszul illeszkedik” a Benford-törvényhez (1-el kezdődik általában), a fenti megfigyelés demonstrálásához tehát tökéletes példa lesz. A példa kedvéért generáltam tízezer darab normális eloszlású véletlen számot 1,75 várható értékkel és 0,1 szórással, majd ennek a halmaznak megvizsgáltam a hatványait. A következő relatív gyakoriságokat kaptam.

kezdő számjegy	Hatványkitevő=						Benford törvény
	1	2	4	8	16	24	
1	99,26%	0,06%	39,02%	36,38%	31,88%	30,28%	30,10%
2	0,74%	43,25%	0,00%	3,70%	15,74%	17,54%	17,61%
3	0,00%	55,95%	0,06%	4,11%	11,15%	12,00%	12,49%
4	0,00%	0,74%	0,43%	6,76%	8,95%	10,34%	9,69%
5	0,00%	0,00%	2,68%	8,82%	8,39%	8,27%	7,92%
6	0,00%	0,00%	8,02%	10,22%	7,52%	6,45%	6,69%
7	0,00%	0,00%	13,53%	11,22%	6,22%	5,80%	5,80%
8	0,00%	0,00%	18,59%	10,05%	5,18%	4,72%	5,12%
9	0,00%	0,00%	17,67%	8,74%	4,97%	4,60%	4,58%

5. Táblázat: A méterben mért emberi testmagasság különböző hatványainak első-számjegy eloszlása

Alkalmazások

A számjegyelemzés fejlődése során hosszú utat járt be. Az első ismert javaslat a Benford-törvény ilyen típusú alkalmazására még 1972-ből származik, egy azóta már híressé vált közgazdász, Hal R. Varian olvasói leveléből⁴⁷. Ő arra hívta fel a figyelmet, hogy az empirikus

⁴⁷ “After all, Benford’s law is just a curious empirical, almost numerological, phenomena; why should it have anything to do with economic forecasting? However, one must admit that if the input data did obey Benford’s law while the output didn’t... well, that would make one feel a trifle uneasy about actually using that out-put.” (Varian [1972])

közgazdasági adatok meglehetősen jó illeszkedése a Benford-törvényre egy lehetséges módja a közgazdasági modellek, az azokból származó előrejelzések jóságának megítéléséhez. Ha ugyanis a modell által generált adatok nem engedelmeskednek a törvénynek, akkor feltételezhetően maga a modell sem jó.

Ez az észrevétel akkoriban nem sok visszhangra talált tudományos körökben (ami nem meglepő egy frissdiplomás szerző olvasói levele esetén). Az első valódi alkalmazások ennek megfelelően a 80-as évek második feléből származnak. Carslaw [1988] a Benford-törvény második számjegyekre vonatkozó verzióját használta annak kimutatására, hogy az új-zélandi cégek éves beszámolóikban következetesen felfelé kerekítik a bevételeiket: a 0 sokkal gyakrabban, a 9 pedig sokkal ritkábban szerepelt második számjegyként, mint az várható lett volna. Ugyanezt az elemzést ismételte meg USA-beli adatokkal Thomas [1989], aki ezen kívül a lefele kerekítésre irányuló tendenciát is talált a jelentésben szereplő veszteségek esetén, és kimutatta, hogy az egy részvényre jutó nyereség (EPS) az 5 cent többszöröse volt, és jóval kevesebbszer végződött 9-re, mint kellett volna.

Magának a jelenség a legfontosabb alkalmazása jelenleg az ellenőrzés, annak is inkább a „forensic” néven emlegetett, „nyomozós” változata, mint például az adóhatósági ellenőrzés⁴⁸. Ez azért lehetséges, mivel Hill [1988] kimutatta, hogy amennyiben az emberek „kitalálnak” számokat, akkor ezek a számok **nem fognak megfelelni** a Benford-törvénynek.

1992 óta egy Nigrini nevű amerikai szakember sok úttörő jellegű alkalmazást fejlesztett ki az auditorok részére. Vannak azonban más alkalmazások a szorosabban vett ellenőrzésen kívül, ám többségükben ezek is az ellenőrzés egy határterületét jelentik bizonyos szempontból. Egy érdekes alkalmazás például a Schraepfer és Wagner [2003] cikkében leírt módszer a hamisított (azaz a kérdezőbiztos által önhatalmúlag kitöltött) interjúk, űrlapok felderítésére statisztikai adatfelvételek esetén.

Egy valóban nem-ellenőrzési alkalmazás lehet talán a processzorok lebegőpontos számításainak optimalizálása. A legfrissebb tudományos kutatások szerint ugyanis a számításokhoz felhasznált inputok eloszlása nem egyenletes, hanem a Benford-törvényt követi. Lehetőség van arra, hogy a későbbi architektúráknál ezt figyelembe véve sebességnövekedést érhessenek el, hasonlóan ahhoz, ahogy a számítógép-írógép

⁴⁸ Érdeklődésemre az APEH-nél elmondták, hogy ők nem alkalmazzák ezt a módszert, szemben például az IRS-sel (USA-beli „APEH”), amely kb. 10 éve használja.

billentyűzetkiosztása (illetve korábban a nyomdászok szedőasztalán a betűk elhelyezkedése) esetében történt.

Az ellenőrzési alkalmazások önmagukban is eléggé szerteágazóak, a következőkben ezért csak a leginkább tipikus elemzési szituációkat ismertem. Miután felsoroltam a főbb alkalmazási területeket, a következő szakaszban áttekintést fogok nyújtani azon statisztikai tesztekéről, amelyeket ezen területek elemzésekor használhatunk. Végül az ezután következő szakasz fogja értékelni a konkrét gyakorlati problémákat, az alkalmazhatóság korlátait.

Maga a számjegyelemzés alapvetően egy kockázatelemzési módszer: a célja megjelölni azon területeket, ahol magasabb a hibás/manipulált adatok előfordulási valószínűsége.

Az első lépés ennek során annak eldöntése, hogy egyáltalán ésszerű-e feltételezni, hogy az ellenőrzés alatt álló sokaságnak Benford-eloszlást kell követnie. A lehetséges eszközök az első, második, harmadik és negyedik számjegyek eloszlásának vizsgálata.

- Az *első számjegyek* eloszlásának vizsgálata általában az első lefolytatandó teszt egy számjegyelemzés során. Ez a tesztet arra lehet és kell használni, hogy a Benford-szerű eloszlás hipotézisének ésszerűségét elfogadjuk/elvessük. Nem alkalmas tehát arra, hogy ez alapján mintát vegyünk.
- A *második/harmadik/negyedik számjegyek* ellenőrzése további bizonyosságot szolgáltathat arról, hogy a sokaság eloszlása valóban leírható a Benford-törvénnyel. Ezek a tesztek sem alkalmasak a mintavétel közvetlen előkészítésére.

A leírt tesztek tehát azt segítenek eldönteni, hogy a sokaságról ésszerű-e feltételezni, hogy a számjegyek eloszlása követi a Benford-törvényt. Ezek a tesztek emiatt úgynevezett *magas szintű tesztek*, eredményükből konkrét mintavételi, kockázati következtetés nem vonható le. Annak érdekében, hogy kijelölhessük a sokaság kockázatos részterületeit, az alábbi tesztek állnak rendelkezésre a második lépcsőben.

- Az *első két számjegy* eloszlásának elemzése (néha együtt az első három, vagy első négy számjegy elemzésével) már közvetetten alkalmazható a minta kijelölésére, de csak közvetett módon. Azon számjegy-kombinációk (pontosabban az ezeknek megfelelő tranzakciók), amelyek gyakorisága szignifikánsan eltér a Benford-törvény szerintitől, majdnem ideális gyanúsítottak a későbbi alaposabb vizsgálatoknál. Mivel azonban a valós számhalmazok csak közelítőleg Benford-szerűek, ezért az első két/három/négy számjegy elemzése sokszor csak vaklarmát eredményez. Mindazonáltal ezek a tesztek már jól jelzik, hogy a számjegyek mely csoportja túl

„divatos”, és hogy hová érdemes a későbbi elemző/alapvető eljárásokat összpontosítani.

- Az *ismétlődés* tesztelése természetes kiegészítője és egyben folytatása is az előbbi tesztnek. A tesztelés módja eltér a szokásostól: egy mutatószámot (NFF – number frequency factor) kell meghatározni minden gyanús részhalmazra⁴⁹. A mutatószám értéke 0 (minden érték ugyanolyan) és 1 (minden érték különböző) közé esik. Közkeletű az a vélekedés, hogy egyes számjegycsoportok túlzott gyakorisága jelezhet rossz szervezést, hibákat, (szándékosan) téves állításokat, és végül, de nem utolsón sorban bűncselekményt (sikkasztás, pénzmosás stb.) is⁵⁰. Minél közelebb van tehát a mutatószám értéke nullához, annál „gyanúsabb” a halmaz.
- A *kerekítési* tesztet akkor érdemes használni, amikor a kerekítés nem fogadható el: az 5, 10, 25, 100, 1000 stb. számok többszöröseinek túlzott gyakorisága jelezhet hanyagságot, de bűncselekményt is. Különösen akkor hasznos ez a teszt, amikor egy céget fel akarnak vásárolni, és a készletek stb. értékét a felvásárlás célpontja szolgáltatja a potenciális „kérőnek”.
- Az *utolsó két számjegy* eloszlásának tesztelése hasonlít a kerekítési teszteléshez. Olyankor érdemes használni, amikor valamilyen okból gyanítható, hogy az adatokat következetesen „kitalálják”, tehát hogy azok teljesen fiktívek. Bizonyítható⁵¹, hogy az n . szignifikáns számjegy, $S_n(x)$ eloszlása n növekedésével tart az egyenleteshez. Az utolsó két számjegy eloszlása nem pont ugyanezt jelenti ugyan, de az empirikus megfigyelések alapján ésszerű feltételezni, hogy az utolsó két számjegy egyenletesen, tehát 1/100 relatív gyakorisággal oszlik el.

Ezek volnának tehát a főbb alkalmazási célterületei a számjegyelemzésnek. A legtöbb modern ellenőrzéstámogató szoftver rendelkezik egy olyan beépített funkcióval (pl. Benford-modul a

⁴⁹ A mutatószám meghatározási módja a következő. A vizsgált adatokat csoportba kell osztani értékük szerint. Az i . értékre jelentse c_i azon tételek számát, amelyek értéke az i . érték. Ezután az egyedül lévő elemeknél változtassuk c értékét 1-ről 0-ra. Ha n db adatot vizsgálunk, a mutatószámot az $NFF = 1 - \frac{\sum c_i^2}{n^2}$ képlet adja.

⁵⁰ Rossz szervezésre példa, amikor a gyakran ismétlődő standard bevételeket (pl. illetékbélyeg eladások az illetékhivatalban) vagy kiadásokat (pl. bérkifizetés) egyenként könyvelik ahelyett, hogy analitikát vezetnének, és a feladásokat könyvelnék. A téves állításra példa az adatok szisztematikus felfelé (vagy lefelé) kerekítése.

⁵¹ Lásd például Hill-Schürger [2005], 4.4

CaseWare IDEA esetén), amely segít ezen elemzések elvégzésében. Az IDEA szoftver ezen moduljából a hátralévő szakaszokban több képernyőmentést is megmutatok.

A következő szakaszban ismertetem azokat a statisztikai teszteket, amelyeket ezekkel a területekkel összefüggésben, de akár önállóan is alkalmazhatunk.

A Benford-törvény érvényesülésének statisztikai tesztelése

Először is meg kell említeni, hogy noha az esetek többségében feltételezzük a sokaság Benford-törvénynek megfelelő viselkedését, ez nem mindig van így. Van ugyanis olyan alkalmazás, nevezetesen a folyamatos auditálás, ahol nem feltétlenül a Benford-törvény a viszonyítási alap, sokkal inkább a régóta ismert szervezetenél a megszokott, szervezetre jellemző számjegyeloszlás, ami nem feltétlenül egyezik meg ezzel.

Az ilyen eltérések a Benford-törvénytől gyanúsak lehetnek akkor, ha az auditor először jár a szervezetenél, de teljesen normálisnak tekinthetőek, ha a korábbi időszakokban is ugyanezt tapasztalta, és nem talált hibát. Ilyen normális, szervezet-specifikus eltérést okozhat például egy kórháznál az, hogy nagyon gyakran szereznek be standard csomagban fix áron injekciós-tűket; ha ezek ára például 5-tel kezdődik, akkor többlet mutatkozik valamelyik 5-tel kezdődő számjegycsoport gyakoriságában. Ez az eltérés azonban a statisztikai tesztek nagy részét nem zavarja, egyszerűen a tapasztalati eloszlást kell viszonyítási alapnak tekinteni a Benford-törvény helyett.

Feltételezzük, hogy bármely eltérés az előírt számjegy-eloszlástól az alábbi két tényező egyikének tudható be:

- Mintavételi hiba, vagy
- Valós hiba, ami lehet például manipuláció, bűncselekmény, illetve tévedés (rögzítési, számolási stb.), esetleg rossz szervezés következménye.

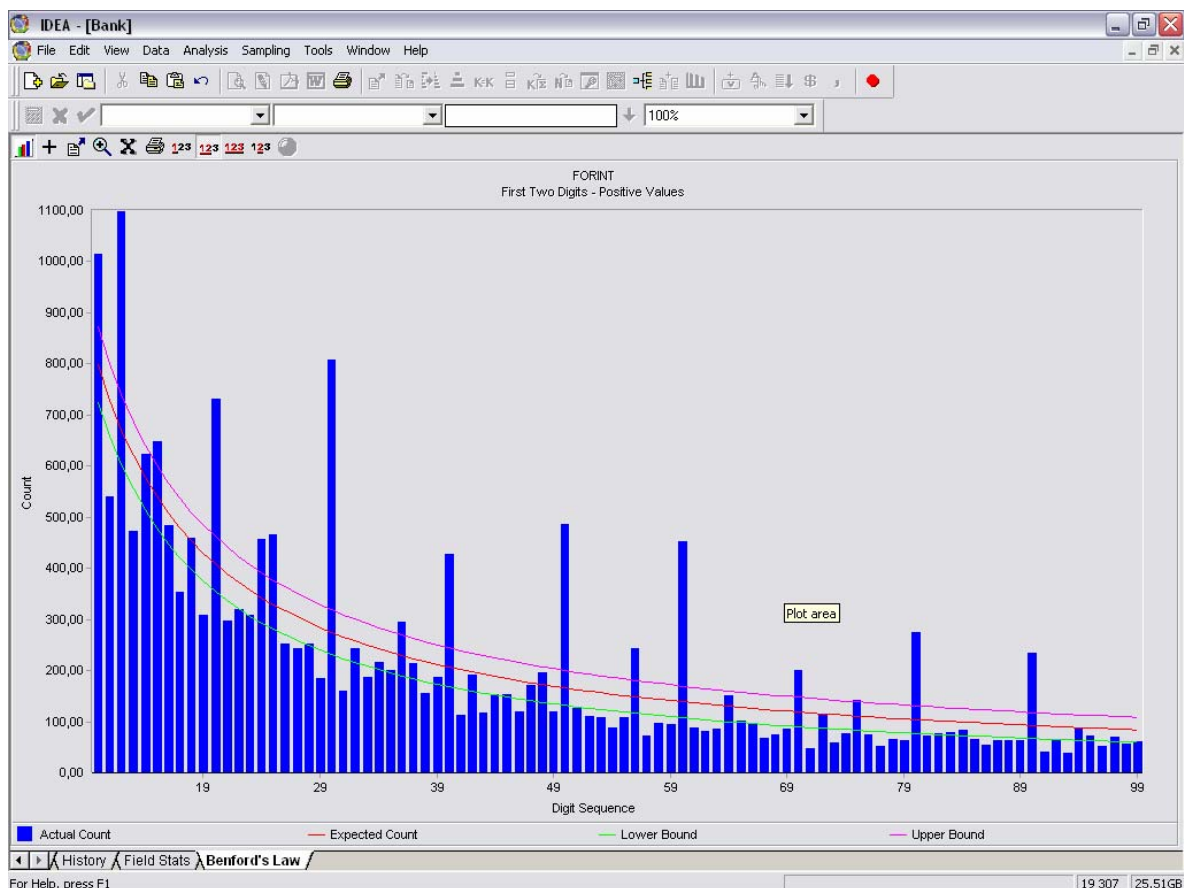
Az eltérés ezen két lehetséges forrása közötti választásra több statisztikai módszer is a rendelkezésünkre áll, ezek legtöbbször alapvetően illeszkedésvizsgálati módszer. A rendelkezésre álló módszerek közül az alábbiakat fogom részletesebben is ismertetni:

1. grafikonok szemrevételezése,
2. Kolmogorov–Szmirnov próba,
3. átlagos abszolút eltérés próba,

4. khi-négyzet próba,
5. z-próba,
6. az elméleti és az empirikus értékek regresszióján alapuló próba,
7. összegző teszt.

Az 1-6 tesztek az empirikus adatok illeszkedését tesztelik a viszonyítási alapként szolgáló eloszláshoz, ami jellemzően a Benford-törvény. A 7. tesztnek más célja van, és a Benford-törvény összeg-invarianciáját használja fel.

A grafikonok szemrevételezése fontos első lépés lehet nem csak a számjegyelemzésben, de a statisztika számos már területén is. Az értékek grafikus ábrázolása gyorsan megmutathatja, merre érdemes továbbmenni, és ezért a legtöbb szoftver támogatja is ezt a funkciót. A számjegyelemzés esetén a leggyakrabban használt grafikon megmutatja az egyes számjegycsoportok gyakoriságát a sokaságban, a Benford törvény alapján várható elméleti gyakoriságokat, és esetleg az erre vonatkozó konfidencia-intervallumokat is. (lásd 4. ábra)



4. ábra: Az adatok vizuális megjelenítése az IDEA szoftverben

A *Kolmogorov-Szmirnov próba* az elméleti és az empirikus eloszlásfüggvények összehasonlítására épül: ezeknek a függvényterben értelmezett maximum norma szerinti távolságát méri. Az egyoldali Kolmogorov-Szmirnov teszt próbafüggvényét tehát a következő képlet adja:

$$D_n(x) = \max |F_n(x) - F(x)|. \quad (2.9)$$

Az első számjegyek esetét alapul véve, az empirikus eloszlásfüggvényt n elemű minta mellett a következőképpen kell meghatározni:

$$F_n(x) = \frac{\#(y : M_{10}(y) \leq x)}{n}, \text{ ahol a } \# \text{ operátor jelöli a „darabszámot”.} \quad (2.10)$$

Ha viszonyítási alapként a Benford-törvény szolgál, akkor az eloszlásfüggvény $\log_{10} x$ (1 és 10 között), más esetben a (2.10) képlettel analóg módon határozható meg. Amíg a viszonyítási alapként szolgáló eloszlás (a nullhipotézis) folytonos, addig a tesztstatisztika eloszlása nem függ a nullhipotézistől. A Kolmogorov-Szmirnov tesztnek a következő korlátai vannak:

- csak folytonos eloszlásokra alkalmazható (a Benford-féle eloszlás folytonos);
- sokkal érzékenyebb az eloszlás közepén, mint a farkán;
- a legsúlyosabb korlát, hogy az eloszlást teljes mértékben specifikálni kell, különben nem érvényesek a kritikus értékek (a Benford-törvény jól specifikált).

A teszt kritikus értékei megtalálhatóak statisztikai táblázatokban és a statisztikai programcsomagokban is.

Az *átlagos abszolút eltérés (MAD) teszt* az empirikus és a viszonyítási alapként szolgáló gyakoriságok átlagos abszolút eltérésén alapul. Formálisan a teszt a következőképpen néz ki az első számjegyekre:

$$MAD = \sum_{i=1}^9 \frac{|p_{\text{empirikus}}^i - p_{\text{benford}}^i|}{9}, \quad (2.11)$$

ami értelemszerűen módosul más esetekre. Néhányan a MAD helyett az MSE mutatót alkalmazzák, de valójában ez sem oldja meg a teszt legnagyobb problémáját, nevezetesen a kritikus értékek hiányát. Ismereteim szerint jelenleg két szubjektív módszer létezik a kritikus értékek meghatározására.

alkalmazás / döntés	Első számjegyek	Második számjegyek	Első két számjegy
jó illeszkedés	< 0.004	< 0.008	< 0.006
elfogadható illeszkedés	0.004 – 0.008	0.008 – 0.012	0.006 – 0.012
Gyenge illeszkedés	0.008 – 0.012	0.012 – 0.016	0.012 – 0.018
nincs illeszkedés	> 0.012	> 0.016	> 0.018

Forrás: Nigrini [2000]

6. Táblázat: A MAD teszt Nigrini által javasolt kritikus értékei

Az első módszer Nigrini nevéhez fűződik: ő a gyakorlati tapasztalatok alapján határozta meg a kritikus értékeket, ezért azokat fenntartással kell kezelni. Az általa javasolt értékek a 6. Táblázatban találhatók.

A másik lehetséges módszer a kritikus értékek meghatározására a Monte Carlo szimuláció (Posch [2004]). Ebben az esetben egy véletlenszerűen generált Benford-halmazt fokozatosan szennyeznek nem-Benford-eloszlású véletlen számokkal. Az egyes fokozatok MAD értékeinek összehasonlítása nagyságrendi becslést ad arra, mennyire lehet szennyezett az eredeti sokaság. Ezt a módszert is érdemes fenntartásokkal kezelni, ugyanis egyrészt az eredmények erősen függenek a nem-Benford változók generálásának módjától, másrészt a szennyezettség növelésével a szimulált MAD varianciája is növekszik.

A *khi-négyzet próba* a statisztika standard módszere az illeszkedésvizsgálatra. Ennek is vannak azonban olyan potenciális hátrányai, amiket figyelembe kell venni az alkalmazás előtt:

- a próbafüggvény eloszlása csak közelítés, tehát kis mintákra az egzakt próbákat (pl. multinomiális) kell előnyben részesíteni;
- a khi-négyzet eloszlás független tagokat feltételez.

Mivel a számjegyelemzést tipikusan nagy, több ezres vagy több tízezres sokaságokon végzik, ezért az első megkötés általában nem számít. Ezzel szemben bizonyos esetekben komolyan megkérdőjelezhető a második, függetlenségi feltétel teljesülése.

A khi-négyzet teszt standard próbafüggvénye:

$$\chi^2 = n \cdot \sum_{i=1}^k \frac{(p_{\text{empirikus}}^i - p_{\text{benford}}^i)^2}{p_{\text{benford}}^i}. \quad (2.12)$$

A megfelelő szignifikanciaszinthez és szabadságfokhoz tartozó kritikus értékek megtalálhatóak a statisztikai táblázatokban és a statisztikai programcsomagokban is.

A *z-próba* tipikus alkalmazása parciális: míg a khi-négyzet próba a teljes adathalmaz illeszkedésének jóságát mutatja, a *z-próba* egy adott kategóriáét. A tesztstatisztikát az ismert módon kapjuk:

$$z = \frac{p_s - p_b \pm \frac{1}{2n}}{\sqrt{\frac{p_b(1-p_b)}{n}}}, \quad (2.13)$$

ahol $1/2n$ a folytonossági korrekció miatt szükséges. A tesztet kétoldali formában alkalmazzák, a kritikus értékek a normális eloszlás táblázatából olvashatóak ki.

	DIGITS	EXPECTED	LOWBOUND	HIGHBOUND	ACTUAL	DIFFERENCE
1	10	799,04	724,42	873,67	1015	-215,96
2	11	729,47	658,17	800,77	541	188,47
3	12	671,05	602,66	739,43	1098	-426,95
4	13	621,29	555,49	687,09	472	149,29
5	14	578,41	514,92	641,90	623	-44,59
6	15	541,07	479,66	602,47	648	-106,93
7	16	508,25	448,74	567,77	484	24,25
8	17	479,19	421,41	536,98	354	125,19
9	18	453,28	397,08	509,48	458	-4,72
10	19	430,02	375,28	484,77	308	122,02
11	20	409,04	355,65	462,43	730	-320,96
12	21	390,01	337,87	442,14	298	92,01
13	22	372,67	321,71	423,63	320	52,67
14	23	356,80	306,94	406,67	309	47,80
15	24	342,24	293,40	391,07	456	-113,76
16	25	328,81	280,94	376,68	466	-137,19
17	26	316,40	269,44	363,36	251	65,40
18	27	304,89	258,80	350,99	242	62,89
19	28	294,19	248,91	339,47	253	41,19
20	29	284,22	239,71	328,72	185	99,22
21	30	274,90	231,13	318,67	807	-532,10
22	31	266,17	223,10	309,24	159	107,17
23	32	257,98	215,58	300,38	242	15,98
24	33	250,28	208,51	292,04	187	63,28
25	34	243,02	201,87	284,17	215	28,02
26	35	236,17	195,61	276,74	200	36,17
27	36	229,70	189,69	269,71	294	-64,30
28	37	223,58	184,10	263,05	213	10,58
29	38	217,77	178,81	256,72	155	62,77
30	39	212,25	173,80	250,71	187	25,25
31	40	207,01	169,03	245,00	428	-220,99
32	41	202,02	164,50	239,55	112	90,02
33	42	197,27	160,19	234,35	191	6,27
34	43	192,74	156,09	229,38	118	74,74
35	44	188,40	152,17	224,64	153	35,40
36	45	184,26	148,43	220,10	154	30,26

5. ábra: A 4. ábra alapadatai az IDEA szoftverben

A 5. ábra mutatja a 4. ábra grafikonjának forrásadatait.⁵² Amint látható, a szoftver automatikusan kiszámolja a gyakoriságokat, és egy magas, 99%-os megbízhatósági szinten

⁵² A Benford törvény gyakorlati alkalmazhatóságának tesztelését az ellenőrzött szervezetektől az ÁSZ ellenőrzések során rutinszerűen bekért főkönyvi adatállományon végeztem.

megadja a kritikus tartományt is (lowbound és highbound oszlopok). Azért 1% és nem 5% a szignifikanciaszint, mert így könnyebben elkerülhetőek a téves riasztások.

Tekintsük például az ábrán a „10” kezdő sorozatot. A teljes 19307 esetből 1015 alkalommal talált ilyet a program, ami 5,2%. Ezzel szemben a Benford-törvény 799 esetet (4,1%) jelez előre, ami szignifikáns eltérés, hiszen 873,67 a felső határ. Egyébként a 4. ábra grafikonján is jól látható, hogy a szoftver automatikusan berajzolta a z-próba kritikus értékeit, ezzel is támogatva a szemrevételezést.

Az elméleti és empirikus értékek regressziója egy szokatlan, de ötletes megoldás. Első lépésként a következő egyenletet becsüljük:

$$p_{\text{empirikus}} = \alpha + \beta \cdot p_{\text{benford}} \quad (2.14)$$

Az $\alpha = 0$ és $\beta = 1$ nullhipotézis (tökéletes illeszkedés) a szokásos eljárásokkal tesztelhető. Vegyük azonban figyelembe, hogy a magas szintű tesztelésre ez a módszer alkalmatlan, ugyanis olyankor csupán 9 vagy 10 megfigyeléssel rendelkezünk. Az alacsonyabb szintű teszteknel már más a helyzet, pl. az első két számjegy együttes eloszlásának tesztelésénél 90 elemű mintával rendelkezünk.

Az *összegző teszt* egy olyan alternatív megközelítés, melynek gyakorlati alkalmazása (Nigrini [1992]) több évvel megelőzte az elméleti megalapozást (Allaart [1997]). Azt a jelenséget teszteli, hogy az adott hosszúságú számjegycsoporttal kezdődő számok összegei megegyeznek-e. A Benford törvénynek eleget tevő halmazokra ugyanis igaz⁵³ az, hogy 1, 2, ... 9 számjegyekkel kezdődő számok összértékei várható értékben megegyeznek, külön-külön a teljes értékösszeg kilenced részét teszik ki. Ugyanez kiterjeszthető az első kettő, három stb. számjegyre is, már ha elég nagy a vizsgált sokaság.

Ez a teszt leginkább arra használható, hogy a nagyságrendi hibákat felfedezzük. Nigrini [2000] idéz egy valós történetet a Wall Street Journalból, mely szerint egy férfi, aki 513 dollár adó-visszatérítésre számított, kapott ehelyett az adóhatóságtól egy fizetési felszólítást 300 000 007,57 dollárról. És ez az eset még kb. 3000 emberrel megtörtént abban az évben. A szokásos számjegyelemzési módszerek ezt (tehát a relatív kis mennyiségű, de nagy értékű „túlszámlázást”) nem tudják kimutatni (3000 eltérés az USA egész lakosságának megfelelő

⁵³ Pontosabban csak akkor igaz, amikor a számok nagyságrendje és a mantisszája közel korrelálatlan, és ez a gyakorlatban általában teljesül is.

sokaságon belül), azonban az összegző teszt igen, mivel a hármassal kezdődő tételek összegét már észrevehetően módosította ez a hiba.

Következtetések, a számjegyelemzés alkalmazási korlátai

Az eddig leírtak alapján a fő kérdés, hogy vajon mennyire reális elvárás, hogy az ellenőrzés során talált adatok számjegyei valóban a Benford-törvénynek megfelelően oszlanak el. Konkrétan a számviteli adatok a gyakorlat szerint általában a törvénynek megfelelően viselkednek, de ez nem jelenti azt, hogy minden olyan főkönyvi adathalmaz, aminek nem ilyen az eloszlása, hibát tartalmaz. Ahogy korábban is írtam, bizonyos szervezet-specifikus tényezők tartósan eltéríthetik a számjegyek eloszlását, de léteznek olyan univerzális tényezők is, amelyek ugyanilyen torzulást eredményezhetnek. A marketingtevékenység például együtt szokott járni a pszichológiai árazás alkalmazásával (198 forint 200 helyett), és senki, aki nyitott szemmel jár nincs meglepve attól, hogy a cégek a lehetőségekhez mérten kozmetikázzák (a jog határain belül) a beszámolóikat. A bank-automatákban történő pénzkivét összege általában ezresre egész szám, így azon cégek banki naplója biztosan torzított lesz, ahol sok az ilyen tranzakció. A különféle előre definiált (pl. jogszabályban előírt) összegű adminisztratív díjak, illetékek szintén okozhatnak hasonlót.

Amennyiben valójában bűncselekmény történt, akkor sem biztos, hogy a számjegyelemzés eredményre fog vezetni. A Benford-törvény akkor sérül, ha a bűncselekményt elkövető úgy ad hozzá, vagy vesz el az eredeti sokaságból tételeket, hogy ezzel megsérti az eredeti eloszlást. Hogy ez a változás szignifikáns legyen, először is szükséges, hogy a módosított adathalmaz mérete elég nagy legyen a teljes adathalmazhoz képest. Ha a cselekmény nem érint tételeket (például mivel azokat nem is rögzítették, mint például a lopásokat), vagy csak nagyon kevés tételt érint, akkor nem várható el, hogy a számjegyelemzés eredményt hozzon. Azon bűncselekmények, amelyek csak olyan adathalmazokat érintenek, amiket a számjegyelemzés nem vizsgál, szintén látókörön kívül maradnak. Példa lehet erre, ha valaki a számlán szereplőtől eltérő címre (például a sajátjára) utalja el a számla ellenértékét. Hasonlóan nehéz felfedezni, ha valaki az összes tételt megszorozza ugyanazzal a számmal: az invariancia miatt ez nem változtat az eloszláson.

Durtschi–Hillison–Pacini [2004] hoz egy jó példát, ami segít annak megítélésében, mekkora a számjegyelemzés alkalmazásának „hozzáadott értéke”. Tegyük fel, hogy a tranzakciók 3 százaléka fiktív, és tegyük fel, hogy az esetek 75%-ában jogosan jelez kockázatot a

sámjegyelemzés (tehát tényleg bűncselekményre, vagy más lényeges hibára akadt). Annak valószínűsége, hogy ténylegesen is hibára akadunk, amikor az elemzés kockázatot jelez:

$$\Pr(\text{találat} \mid \text{jelzés}) = \frac{0.75 \cdot 0.03}{0.75 \cdot 0.03 + 0.25 \cdot 0.97} = 0.085 \approx 9\% \quad (2.15)$$

Ez a 9% majdnem háromszor annyi, mint amit egyszerű mintavétel esetén kaptunk volna, azonban még mindig nagyon alacsony, 10% alatti valószínűség. A számjegyelemzés tehát nem csodafegyver, de ennek ellenére hasznos kiegészítő az ellenőrök fegyvertárában.

Végezetül tekintsük a 7. Táblázatot, ami röviden összefoglalja, hogy az ellenőrzés során mikor célszerű és mikor nem célszerű számjegyelemzéssel próbálkozni.

Mikor lehet hasznos a számjegyelemzés	Példák
sámok halmaza, ahol az értékek két különböző adatsor valamilyen matematikai kombinációjából adódnak	vevők tételei (eladott mennyiség * ár), szállítók tételei (vásárolt mennyiség * ár)
tranzakció szintű adatok – nem kell mintát venni	kiadások, bevételek, költségek
nagy vegyes adathalmazok – minél nagyobb, annál jobb	teljes év tranzakciói
amikor az átlag nagyobb a mediánál, és pozitív a ferdeség	a legtöbb számveteli adat
Mikor alkalmatlan a számjegyelemzés	Példák
mesterségesen kiosztott számok	csekkek sorszáma, számlaszámok, irányítószám
az emberi gondolkodás által befolyásolt számok	pszichológiai árak (pl. 198 Ft), ATM pénzkivét
szervezet-specifikus tételeket tartalmazó számlák	a 100 dolláros visszatérítések számlája
beépített maximummal és/vagy minimummal rendelkező számlák	olyan eszközoldali számla, ahová csak bizonyos érték felett könyvelnek tételeket
nincs nyoma a tranzakciónak	lopás, korrupció stb.

Forrás: Durtschi et al. [2004]

7. Táblázat: Mikor használható és nem használható a számjegyelemzés

III. rész: Aránybecslésre visszavezethető ellenőrzési feladatok

Az ellenőrzés során számos esetben szembesül azzal egy auditor, hogy meg kell határoznia egy sokaságon belül bizonyos tulajdonságú elemek arányát. Ez leggyakrabban a kontrollal védett tranzakciókon belül azon tranzakciók arányát jelenti, ahol a kontroll nem működött, de jelenthet egyszerű pontbecslést is, tipikusan egy teljesítmény-ellenőrzés esetén (például fel kell mérni azt, hogy egy adott új közszolgáltatásról az érintettek mekkora része hallott).

Az I. részben már szó volt róla, hogy szervezet kontrolljainak megítéléséhez egyrészt meg kell vizsgálni, hogy a kontrollok elméletileg képesek-e ellátni feladatukat, másrészt pedig, hogy végrehajtják-e az előírt kontrollokat. További kérdés szokott lenni, hogy ezeket a kontrollokat dokumentálták-e, és végül, hogy a kontroll végrehajtójának volt-e megfelelő felhatalmazása a feladat elvégzésére.

Ha a kontroll csak nagyon ritkán esik ki, kicsi a valószínűsége, hogy ugyanakkor szükség is lett volna rá. Minél gyakrabban esik ki egy kontroll, annál inkább megnő ez a valószínűség. A könnyebb érthetőség kedvéért ezt egy egyszerű példával is illusztrálom. A példa szándékosan nem gazdálkodási jellegű kontroll, mutatva a gondolat univerzális jellegét. Legyen tehát a vizsgált kontroll a beléptetés ellenőrzése, ahol 4 órás váltásban 1-1 portás látja el a feladatot. Ha a portás a 4 órás műszak alatt egyszer belenéz az újságjába, az talán megengedhető. Ha elolvassa a teljes vezércikket, vagy elmegy egyszer a mellékhelyiségbe, talán még az is. De az már általában nem, ha az idő nagy részében a cigarettázó helyiségben tartózkodik (a portától távol). Persze minden attól függ, hogy egy poros új-mexikói motel, vagy egy kémelhárító szervezet portájáról van-e szó.

A kontrollok kiesése tehát egy bizonyos határig tolerálható, egy bizonyos határ felett azonban nem. Annak megállapítása, hogy hol húzódik ez a határ, például az első részben leírt költség-haszon elv alapján történhet, de mindenképpen szakmai döntés eredménye, aminek elemzése túlmutat a disszertáció keretein. A jelen disszertáció szempontjából tehát a *hibák (kiesések) tolerálható aránya* a vizsgált kontrollra nézve *adottnak* tekinthető.

Szintén külön értekezést érne meg annak tisztázása, mikor tekintjük megvalósultnak a kontrollt, és mikor nem. A fenti esetnél maradva, elképzelhető, hogy például az újságolvasás egyáltalán nem akadályozza a portást a feladatellátásban, mert az ajtónyitás csak a csöngetés

után, a portáról lehetséges. De elképzelhető olyan helyzet is (és ez a gyakoribb), amikor az újságolvasás már lehetővé teszi az illetéktelen belépést. Mi most feltételezzük, hogy a kontroll megvalósulása vagy kiesése *egyértelműen eldönthető* minden konkrét esetben.

A gyakorlatban további problémát jelent annak megállapítása, hogy *mely tételek*, konkrét tranzakciók *esetében kellett volna működnie a kontrollnak*. Ez az alkalmazóknak azért fontos, mert ez a halmaz képezi azt a sokaságot, amiből a tesztelés során a mintát kell majd venni. Ám ennek a meghatározási módja is túlmutat a disszertáció keretein.

A kontrollteszteléssel az auditor célja az, hogy a megvizsgált tranzakciók, mint minta alapján eldöntse, megbízhatóan működik-e a kontroll, tehát a kiesések aránya tolerálható-e. Ez a cél lehet „öncél”, amikor például az ellenőrzés célja magának a kontrollrendszer megfelelő működésének, mint állításnak az igazolása, de lehet eszköz is, amikor el szeretné dönteni, az bízhat-e a vizsgált kontroll működésében, és megelégedhet kevesebb alapvető teszteléssel, vagy el kell vetnie a kontrollt, és a nagyobb kockázat miatt több erőforrást kell erre a területre csoportosítani.

A kontrolltesztelés időszerezése: SOx 404

Az első részben már említett Sarbanes-Oxley Act 404-es szakasza (Assessment of internal control) előírja a menedzsmentnek és a külső könyvvizsgálónak, hogy jelentésükben térjenek ki a beszámoló készítésével kapcsolatos kontrollok⁵⁴ megfelelőségére. Minden kétséget kizáróan a végrehajtás szempontjából ez a jogszabály leginkább „költséges” része, hiszen a releváns manuális és automatikus pénzügyi kontrollok dokumentálása és tesztelése rendkívüli erőfeszítéseket igényel az érintettektől⁵⁵.

A SOx által létrehozott Public Company Accounting Oversight Board (PCAOB) által kiadott ellenőrzési standard a következő követelményeket támasztja a belső kontrollok felmérésével kapcsolatban a külső könyvvizsgálók felé:

⁵⁴ ICFR – internal control over financial reporting

⁵⁵ Ezek a költségek azonban nem hiábavalóak. Ashbaugh-Skaife et al. [2007] tanulmánya például kimutatta, hogy azon cégek tőkeköltsége, amelyek jelentősen javították a kontrollrendszerüket a SOx-szal összhangban, 50-150 bázisponttal csökkent. Az IIA (belső ellenőrök nemzetközi szervezete) felmérései pedig arra mutatnak rá, hogy a SOx következtében sokat javult az éves beszámolók megbízhatósága.

- fel kell mérni mind a kontrollok kialakítását („design”) mind a működésük megbízhatóságát („operating effectiveness”);
- elemezni kell a folyamatokat (az IT folyamatokat is ideértve) és azonosítani kell a folyamatokban lévő kockázatos pontokat;
- értékelni kell a kialakított belső kontrolloknak az első COSO keretrendszerrel való összhangját;
- fel kell mérni a csalás kockázatát;
- értékelni kell a csalás ellen létrehozott kontrollokat, ideértve azon átfedő kontrollokat is, amelyek a kontrollok menedzsment általi áthágását hivatottak meggátolni;
- értékelni kell a zárási folyamatba épített kontrollokat;
- a felmérés mélységének összhangban kell lennie a vizsgált szervezet méretével és a folyamatok összetettségével;
- a menedzsment munkájára való támaszkodáskor figyelembe kell venni olyan tényezőket is, mint az érintettek hozzáértése és objektivitása, valamint figyelembe kell venni az ehhez kapcsolódó kockázatokat;
- értékelni kell a vagyon megőrzése céljából működtetett kontrollokat; és végül
- mindezek alapján véleményt kell mondani a beszámolással kapcsolatos kontrollok megfelelőségéről.

A menedzsment felé ehhez hasonló követelményeket támaszt az SEC útmutatója.

A kontrolltesztelés alkalmazása az Állami Számvevőszék munkájában

Az Állami Számvevőszék ellenőrzési kézikönyve (a „zöld könyv”) alapvetően két helyen hivatkozik a kontrollok tesztelésre: a financial audit típusú ellenőrzéseket előkészítő kockázatelemzésnél és az átfogó ellenőrzéseknél. Noha a kockázatelemzés végrehajtásával jelentős ellenőrzési erőforrásokat lehetne megtakarítani, az ÁSZ jelenlegi gyakorlatában a kockázatelemzés és ennek keretében a kontrolltesztelés is inkább formális jellegű, ugyanis az alapvető eljárások mintaméretét a kockázatelemzés eredménye a gyakorlatban nem befolyásolja. A másik lehetséges alkalmazás az átfogó ellenőrzés, amiről a „zöld könyv” így ír: „az átfogó ellenőrzés a szervezetek működését, projektek, tevékenységek végrehajtását

szolgáló belső irányítási, szabályozási (kontroll) és ellenőrzési rendszerre irányul”. A gyakorlatban egészen a 2006. év végéig ennek a kritériumnak sem egyetlen átfogó ellenőrzés, sem a kapcsolódó módszertanok, segédletek, ellenőrzési programok nem feleltek meg.

2007. elejétől való alkalmazási hatállyal kidolgozásra került az önkormányzati területre egy új ellenőrzési program, amely úttörő módon megcélozta a „zöld könyv” definíciójának való megfelelést (az államháztartás központi szintjén folytatott átfogó ellenőrzések továbbra is a régi megközelítést alapján folynak). Ezzel párhuzamosan az átfogó ellenőrzések kidolgozás alatt álló módszertana is kiemelt hangsúlyt fektet a különféle kontrollok ellenőrzésére, tesztelésére.

Az újonnan kidolgozott önkormányzati átfogó ellenőrzések célja annak értékelése, hogy az önkormányzat:

- a pénzügyi egyensúlyt a költségvetésében és annak teljesítése során milyen módon biztosította, a teljesített bevételek és kiadások egyes évek közötti jelentős eltérése feladatváltozáshoz kapcsolódott-e;
- felkészült-e a szabályozottság és a szervezettség terén az európai uniós források igénylésére és felhasználására, továbbá az e-közigazgatás bevezetése miatti szervezetszerűsítési feladatokra;
- *kialakította-e a külső és a belső feltételeknek megfelelően a gazdálkodás belső kontrollrendszerét, továbbá a költségvetés tervezési, végrehajtási és zárszámadási feladatok szabályszerű ellátásához hozzájárult-e a folyamatba épített, előzetes és utólagos vezetői ellenőrzés, valamint a belső ellenőrzés;*
- megfelelően hasznosították-e a korábbi számvevőszéki ellenőrzések megállapításait, szabályszerűségi és célszerűségi javaslatait.

A dőlt betűs rész tartalmazza a kontrollteszteléssel összefüggő célt. Ezen belül a releváns részszerzőpontok:

- a költségvetési tervezés és a zárszámadás készítés folyamatában teljesítették-e az előírt belső kontroll feladatokat;
- megfelelően működtek-e a gazdálkodás folyamatában a szakmai teljesítés igazolás és az utalvány ellenjegyzés kontrolljai az általánostól jellemzően eltérő, egyedi eljárás igénylő gazdasági eseményekkel (az állományba nem tartozók megbízási díjaival, a

karbantartási, kisjavítási szolgáltatásokkal, továbbá a gépek, berendezések felszerelések beszerzésével) kapcsolatos kifizetések során;

- segítették-e a pénzügyi-számviteli feladatok ellátását az informatikai rendszerek, azok biztonságos, dokumentált működési feltételeit kialakították-e.

Az ellenőrzés végrehajtásának támogatására, az ellenőrzést végzők munkájának megkönnyítése érdekében részletes útmutató készült az összes ellenőrzési kérdéshez, ami tartalmazza a lefolytatandó ellenőrzési eljárások leírását. A kontrollok tesztelésére kidolgozott módszert és a módszer alkalmazásának eredményeit a szekvenciális eljárásokról szóló részben ismertetem.

Pontbecslések

Noha a sokasági arány pontbecslése az ellenőrzési munkában alapvetően a közvélemény-kutatáshoz hasonló alkalmazásokban (pl. elégedettség-mérés) szükséges, az értekezésben mégis kitérek rá röviden. Ezt az indokolja, hogy a bayesi pontbecslések relevanciája – pontosan az előzetes információk felhasználhatósága miatt – az egyéb, nem ellenőrzési alkalmazásokhoz képest magas, ám ennek ellenére a hazai ellenőrzési szakmában még nem nagyon ismerték fel a bayesi becslések gyakorlati hasznát. Ebben a pontban tehát röviden áttekintek pár egyszerűbb pontbecslési módszert a sokasági arány meghatározására, nem térek ki azonban a mintavétel reprezentativitásával, vagy a non-response témával kapcsolatos problémákra, ugyanis ezeknek – az amúgy érdekes – részletkérdéseknek az ismertetése álláspontom szerint elterelné a figyelmet a fontosabb kérdésekről.

A sokasági arány ($P = \frac{M}{N}$) pontbecslése egyszerű véletlen mintavétel használatakor a megfelelő mintabeli értékkel ($\hat{P} = p = \frac{m}{n}$) történik. A mintabeli érték torzítatlanul becsüli a sokasági értéket:

$$E(\hat{P}) = E\left(\frac{m}{n}\right) = \frac{1}{n} E(m) = \frac{1}{n} \cdot n \cdot \frac{M}{N} = \frac{M}{N}, \quad (3.1)$$

ugyanis m jelen feltételek mellett hipergeometriai eloszlást követ, és a hipergeometriai eloszlás várható értéke $\frac{nM}{N}$.

Alternatív megoldást kínál a pontbecslésre a bayesi megközelítés. Ehhez priorként szolgál, ha a sokasági hibaszámról feltesszük, hogy binomiális eloszlású (N, P_0) paraméterekkel, ahol P_0

a korábbi ellenőrzési tapasztalatokból ismert, rendszersajátosságot tükröző stabil paraméter. A minta hibaszámáról tudjuk, hogy adott sokasági hibaarány mellett N , M , n paraméterű hipergeometriai eloszlást követ, tehát így már meghatározható $\frac{M}{N}$ (pontosabban M) posterior eloszlása is. A posterior eloszlás ismeretében többféle elven is (például a posterior eloszlás várható értéke, mediánja, módusza stb.) készíthető pontbecslés, de ezek részletezését terjedelmi okok miatt mellőzzük. Céljainknak megfelel, ha becsült értéként a posterior eloszlás móduszát adjuk meg, tehát azt a sokasági hibaarányt, amely mellett legvalószínűbb a kapott minta előfordulása. Választásunkat egyrészt indokolja a módusz viszonylag könnyű meghatározhatósága, másrészt pedig vonzóvá teszi a „maximum-likelihood”-szerű interpretálhatóság is. Megjegyezném, hogy hátránya is van a módusz választásának, ugyanis az amúgy „jól viselkedő” (értsd: „haranggörbe-szerű”) eloszlásoknak diszkrét esetben két – amúgy szomszédos – módusza is lehet. Az egyértelműség kedvéért mi következetesen a nagyobb móduszt fogjuk használni.

Gyakorlatban a módusz meghatározásához minden $M \in [m; N - n + m]$ paraméterre ki kell számolni a posterior valószínűséget, és meg kell keresni azt az M -et, ahol ez felveszi a maximumát:

$$f_m(M) = \frac{\binom{N}{M} P_0^M (1-P_0)^{N-M} \binom{M}{m} \binom{N-M}{n-m}}{\sum_{i \in [m; N-n+m]} \left\{ \binom{N}{i} P_0^i (1-P_0)^{N-i} \binom{i}{m} \binom{N-i}{n-m} \right\}} \quad (3.2)$$

Ismert N , n és m mellett ez a számítás nagyon könnyen elvégezhető, például az Excel felhasználásával. Ennek ellenére levezethető egy egyszerű képlet, melynek eredménye interpretálható úgy, mint a mintabeli hibaarány, és a mintán kívüli rész speciálisan számított prior hibaarányának a súlyozott átlaga.

Algebrai átalakításokkal ugyanis a (3.2) posterior valószínűség-eloszlást jóval egyszerűbb alakra hozhatjuk. Először bontsuk ki a binomiális együtthatókat:

$$f_m(M) = \frac{\frac{N!}{M!(N-M)!} P_0^M (1-P_0)^{N-M} \frac{M!}{m!(M-m)!} \frac{(N-M)!}{(n-m)!(N-M-n+m)!}}{\sum_{i \in [m; N-n+m]} \left\{ \frac{N!}{i!(N-i)!} P_0^i (1-P_0)^{N-i} \frac{i!}{m!(i-m)!} \frac{(N-i)!}{(n-m)!(N-i-n+m)!} \right\}} \quad (3.3)$$

Látható, hogy egyszerűsíthetünk elég sok taggal, ami után ez marad:

$$f_m(M) = \frac{P_0^{M-m} (1-P_0)^{N-M-n+m} \frac{1}{(M-m)!(N-M-n+m)!}}{\sum_{i \in [m; N-n+m]} \left\{ P_0^{i-m} (1-P_0)^{N-i-n+m} \frac{1}{(i-m)!(N-i-n+m)!} \right\}} \quad (3.4)$$

Ha ezt bővítjük $(N-n)!$ -sal, és alkalmazzuk a binomiális tételt, akkor a következő nagyon egyszerű formulához jutunk:

$$f_m(M) = \binom{N-n}{M-m} P_0^{M-m} (1-P_0)^{N-n-(M-m)} \quad (3.5)$$

Ezt a formulát megvizsgálva látszik, hogy a gyakoriságfüggvények bármely m esetén egymás „x-tengely” (a független változó, jelen esetben M) mentén való eltoltsai, ezért elégséges csak az $m=0$ esetre elvégezni a számítást, a többi értéket az $f_x(M) = f_0(M-x)$ azonosság felhasználásával kapjuk. Ha $m=0$, a következő, egyszerűsített formulát kapjuk:

$$f(M) = \binom{N-n}{M} P_0^M (1-P_0)^{N-n-M} \quad (3.6)$$

A szélsőérték megkereséséhez a deriválás diszkrét analógiáját alkalmazhatjuk. M egységnyi növekedésekor (3.6) a következőképpen változik meg:

$$\frac{f(M+1)}{f(M)} = \frac{N-n-M}{M+1} \cdot \frac{P_0}{1-P_0} \quad (3.7)$$

Ez a függvény M -ben monoton csökken, tehát a (legnagyobb) módusz annál a legkisebb M -nél lesz, ahol a függvény értéke egynél kisebb (több módusz esetén a nem maximális móduszokra a monotonitás miatt a hányados értéke 1). Ezt a feltételt felírva a $m=0$ -ra a módusz helye:

$$M_{m=0} = \lceil (N-n+1)P_0 - 1 \rceil, \quad (3.8)$$

ahol $\lceil \rceil$ jelöli a felfelé kerekítés műveletét. Ez felhasználva kapjuk a végleges bayesi pontbecslést, ami valójában egy súlyozott átlag:

$$\hat{P} = \frac{M_{m=0} + m}{N} = \frac{N-n}{N} \cdot \frac{M_{m=0}}{N-n} + \frac{n}{N} \cdot \frac{m}{n}. \quad (3.9)$$

A pontbecsléssel kapcsolatos utolsó érdekesség egy brit biológus-genetikus, Haldane [1945] által a *Biometrica*-ban leírt ötlet, amely később „inverz mintavétel” néven vált ismertté. Ő azt javasolta, hogy ritka jelenségek arányának meghatározásakor ne a mintaméret legyen fix, hanem a talált eltérések száma. Ez a gyakorlatba átültetve azt jelenti, hogy egészen addig történik *visszatevéssel* a mintavétel, amíg a szükséges számú eltérést fel nem fedezzük. Ilyen

mintavételi terv mellett a sokasági arány torzítatlan pontbecslése – az intuícióval ellentétben – nem a mintabeli arány, hanem helyett $\hat{P} = \frac{m-1}{n-1}$, ahol m az előre rögzített eltérések száma, és n a mintaelemszám.

Intervallumbecslés, hipotézisvizsgálat (egylépcsős mintavétel)

A sokasági arány meghatározására irányuló ellenőrzési alkalmazások nagyobb része tesztelés, tehát intervallumbecsléssel, illetve hipotézisvizsgálattal kerül kivitelezésre. A hagyományos módszer, ami a gyakorlati alkalmazások közel 100%-át lefedi, az egylépcsős mintavétel. Egy későbbi fejezetben kerül sor egy másik megközelítés, a szekvenciális mintavétel ismertetésére. Ez utóbbi lehetővé teszi a mintaméret csökkentését, és ezáltal az ellenőrzésre fordított erőforrások csökkentését.

A hibamentes mintából levonható következtetések

A kontrolltesztek egy része gyorseszt abban az értelemben, hogy alkalmazásukkor csak nagyon kis mintát vesznek, és ha abban *nem találunk hibát*, akkor a későbbiekben támaszkodnak a kontrollra, egyébként nem. Ezt a módszert, tehát hogy csak a hibamentes minta elfogadható, hívják sokszor discovery samplingnek, azaz felderítési mintavételezésnek. Az ilyen eljárásnak az a hátránya, hogy extrém nagy a másodfajú hiba elkövetési valószínűsége.

A felderítési mintavételezés tervezésénél a mintaméret meghatározása a feladat: olyan mintaméret meghatározása, ami mellett a hibamentes minta elegendő bizonyosságot szolgáltat a kontroll tolerálható hibaarányával (vagy kisebbel) való működésére.

Ám nem csak ebben a helyzetben fordulhat azonban elő, hogy a mintavételt követően a mintában nem találunk hibát. Nulla hiba esetén az egzakt becsléshez szükséges feltételes valószínűségek a következők:

$M = \dots$	0	1	2	...	i
$\Pr(m = 0 M)$	1	$\frac{N-n}{N}$	$\frac{(N-n)(N-n-1)}{N(N-1)}$		$\prod_i \frac{N-n-i+1}{N-i+1}$

8. Táblázat: A nulla mintabeli hiba valószínűsége a sokasági hiba függvényében

Ebben az esetben a becsült felső határ a legkisebb olyan M , amire a valószínűség már éppen kisebb, mint a szignifikancia-szint. Általánosságban ez a feladat nem oldható meg számítógép használata nélkül, ezért hasznos hüvelykujjszabályként két, zsebszámológéppel is kiszámítható közelítést szoktak használni. Ezek azon alapulnak, hogy nagyobb sokaságokra a hipergeometria eloszlás közelíthető binomiálissal, illetve alacsony hibaarányok mellett a binomiális eloszlás közelíthető a könnyen kezelhető Poisson eloszlással.

A binomiális eloszlás esetén a felső határ oszlopában a következő képlet szerepelne:

$$\Pr(m = 0 | M = Np) = \alpha = (1 - p)^n, \quad (3.10)$$

amiből a megfelelő megbízhatóságú becslés:

$$p_{1-\alpha}(0, n) = 1 - \sqrt[n]{\alpha} \quad (3.11)$$

Hibamentes 30 elemű minta és 95% megbízhatóság esetén a becsült felső határ a sokasági hibaarányra 9,5%, 100 elemű mintánál 2,95%.

Még egyszerűbb a Poisson eloszlás használata. Erre az eloszlásra igaz⁵⁶, hogy

$$p_{1-\alpha}(m) \approx \frac{\lambda_{1-\alpha}(m)}{n}, \text{ ahol } \lambda_{1-\alpha}(m) \text{ a Poisson eloszlás paraméterére vonatkozó } 100 \cdot (1 - \alpha) \% \text{-}$$

os felső határ m megfigyelt hiba mellett. Ennek a felső határnak az értékei megtalálhatóak a legtöbb ellenőrzési szakkönyvben („Poisson-táblázat”), de röviden ismertetem a meghatározás módját.

Annak valószínűsége, hogy egy λ paraméterű Poisson-eloszlásból származó ξ valószínűségi változó 0 értéket vegyen fel:

$$\Pr(\xi = 0) = \frac{\lambda^0}{0!} e^{-\lambda} = e^{-\lambda} \quad (3.12)$$

Az intervallumbecslés és a hipotézisvizsgálat dualitása miatt a felső határ becslés megadása megegyezik az alsó elutasítási tartomány megadásával: meg kell találni azt a λ -t, ahol ez a valószínűség megegyezik a szignifikanciaszinttel. Ez 5% esetén $\lambda = 3$.

Ha tehát 95%-os megbízhatóságot szeretnénk, $\lambda_{0,95}(0) = 3$ helyettesítéssel egy nagyon jó hüvelykujjszabály-szerű képletet kapunk a 0 talált hibás esetre. Ha a minta elemszáma elég nagy, de a sokasághoz képest elhanyagolható, a $3/n$ közelítés minden gyakorlati szempontból

⁵⁶ A binomiális és a Poisson eloszlások itt említett kapcsolata megtalálható a legtöbb alapozó valószínűség-számítás könyvben, lásd pl. Denkinger [1990] tankönyvét.

elégseges. 30 elemű mintánál ez a képlet 10%-ot ad, 100 elemű mintánál 3%-ot, nagyságrendileg mindkettő megegyezik a binomiális esettel.

Ha nagyon alacsony ($p \ll \alpha$) a korábbi tapasztalatok alapján várható hibaarány, megfigyelt hiba nélküli minta mellett felmerülhet olyan igény, hogy bizonyos megbízhatósági szinten kijelentsük: a vizsgált sokaságban **nincs** hiba. Adott szuperpopulációs hibaarány (p) mellett annak a valószínűsége, hogy a teljes sokaságban nincs hiba: $(1-p)^{N-n}$, ugyanis ez megegyezik annak valószínűségével, hogy a p hibaarányú végtelen sokaságból a véletlenszerűen kiválasztott $N-n$ tétel egyike sem hibás. Ezért tehát ahhoz, hogy $1-\alpha$ megbízhatósági szinten – a minta alapján – kijelenthessük, hogy a sokaságban nincs hiba, a szükséges mintanagyság:

$$n = N - \frac{\ln(1-\alpha)}{\ln(1-p)}, \quad (3.13)$$

ahol $\ln$ a természetes alapú logaritmust jelöli. A gyakorlati alkalmazás szempontjából ez a módszer csak akkor hasznos, ha N alacsony ugyan (<25), de az egy tranzakcióra eső ellenőrzési költség nagyon magas (például külső szakértőt kell igénybe venni, vagy túlzott időráfordítást jelentene a pótlólagos 5-10 tranzakció ellenőrzése).

A mintavételi statisztikához kapcsolódó becslések

A feltételeket illetően ebben az alfejezetben a vizsgálat tárgyát képező P illetve M sokasági paramétert nem tekintjük valószínűségi változónak. Ismertetésre kerül három eljárás: az első (M1) a standard nagymintás becslés kis mintákon alkalmazva, a második egy ennél pontosabb, de továbbra is közelítő megoldás (M2) és végül az egzakt becslés (M3).

Nagymintás alapeset (M1)

Kiinduló pontunk a tankönyvekben is részletesen tárgyalt becslés, amely esetén a visszatevésees nagy minta esetén a normális eloszlással való közelítés jogosnak tűnik. Ez az eset nem felel meg az ellenőrzési mintavétel követelményeinek, de itt, mint kiinduló pontot, mint etalont tekintjük.

Az eljárás a következő: „*végtelen*” (legalább több ezres nagyságrendű) sokaságból, vagy pedig kisebb sokaságból, de *visszatevéssel* történő mintavétel esetén a mintában található minősített elemek száma binomiális eloszlást követ. A mintabeli hibaarányról (p)

bizonyítható, hogy szintén binomiális eloszlású, továbbá $E(p)=P$ és $Var(p)=\frac{P(1-P)}{n}$, ahol P jelöli a sokasági hibaarányt.

Ha $\min\{np, n(1-p)\} \geq 10$ (tehát legalább 10 hibás és 10 nem hibás elemet találtunk a mintában), akkor a minta hibaarányának transzformáltja közelítőleg standard normális eloszlást követ, amiből az intervallumbecslés:

$$p \pm z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{p \cdot (1-p)}{n}},$$

ahol $z_{1-\frac{\alpha}{2}}$ a standard normális eloszlás megfelelő kvantilise.

Az ellenőrzési gyakorlatban a gazdasági folyamatok jellegéből adódóan 5-10%-nál nagyobb arányú hibát már komolyabb kockázati tényezőként szokás figyelembe venni. Ha ezek alapján megvizsgáljuk ennek a módszernek az alkalmazási lehetőségeit, kiderül, hogy legalább többszáz elemű minta szükséges az $np \geq 10$ feltétel teljesítéséhez, ami a gyakorlati alkalmazások esetében általában nem valósul meg, így a közelítő eljárás alkalmazása torzítást visz az intervallumbecslésbe. További torzításra ad okot a modell kezdeti feltevése, azaz a végtelen sokaságból, vagy pedig visszatevéssel történő mintavétel. A visszatevés nélküli mintavétel esetén a mintabeli hibaarány varianciáját csökkenti egy $\frac{N-n}{N-1} < 1$ szorzó, ezért tehát ennek a feltevésnek a „megszegése” elvileg „jó” irányú torzítást okoz. A teljes torzítás mértékéről analitikusan nehéz pontosan nyilatkozni, de ha összevetjük ezt a becslést a később bemutatásra kerülő M2 becselőfüggvényünkkel, látható, hogy a mintaméret növekedésével (és a kiválasztási arány nullához tartásával) a két becslés is konvergál egymáshoz. Az egyes becslési eljárások pontos torzítási mértékét meghatároztam, ennek részleteiről bővebben később számolok be.

Hipergeometriai eloszlás – normális közelítés (M2)

A közelítő eljárások bevezetésére annak idején főként azért került sor, mivel sokáig nem álltak rendelkezésre táblázatok a hipergeometriai eloszláshoz, így valamilyen folytonos eloszlással helyettesítették a diszkrét eloszlást. Mivel a becslés végső célja a hibaarány meghatározása, a közelítő eljárásokban m és M szerepét általában $\frac{m}{n}$ és $\frac{M}{N}$ veszi át, és ezen

kívül legtöbbször figyelmen kívül hagyják az előbbi hányadosok diszkrét jellegéből adódó sajátosságokat is.

Jelen esetben közelítsük a feltételes eloszlást olyan normális eloszlással, melynek várható értéke és varianciája megegyezik a megfelelő hipergeometriai eloszlásával (Prékopa [1972]). A normális eloszlás kvantilis értékeit használva adott P sokasági hibaaránynál a releváns mintabeli hibaarányok (x) halmaza:

$$x \in \left[P - z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{P \cdot (1-P)}{n} \frac{N-n}{N-1}}; P + z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{P \cdot (1-P)}{n} \frac{N-n}{N-1}} \right]. \quad (3.14)$$

A konstrukció során azzal a feltevessel élünk, hogy bármely x mintabeli hibaarány konfidencia intervallum végpontjához tartozó sokasági hibaarány fent definiált halmazának x határpontja. Mivel folytonos közelítést alkalmazunk, ez a feltevés tartható, továbbá látszik, hogy ebben az esetben a konfidencia intervallum „felső” végpontja a halmaz alsó határa lesz, és fordítva. Írjuk fel ezt egyenletrendszer formájában:

$$x = P_f - z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{P_f \cdot (1-P_f)}{n} \frac{N-n}{N-1}} \quad (3.15)$$

$$x = P_a + z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{P_a \cdot (1-P_a)}{n} \frac{N-n}{N-1}}. \quad (3.16)$$

A feladatunk az, hogy ebből az egyenletrendszerből kifejezzük P_a -t és P_f -et. Könnyen látható, hogy átrendezés és négyzetre emelés után a két egyenlet ugyanúgy fog kinézni, és mivel szintén könnyen látható, hogy P_a -ban, illetve P_f -ben másodfokú egyenletet kapunk, P_a lesz a kisebbik, és P_f a nagyobbik gyök.

Az egyenlet megoldásához először is végezzük el a $c = z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{N-n}{N-1}}$ helyettesítést, ami után

egyenletünk az $x = p \pm c \cdot \sqrt{\frac{p \cdot (1-p)}{n}}$ formát ölti.

Átrendezve és négyzetre emelve:

$$nx^2 - 2npx + np^2 = c^2 \cdot p \cdot (1-p) \quad (3.17)$$

Ismét átrendezve:

$$(n + c^2)p^2 - (2nx + c^2)p + nx^2 = 0 \quad (3.18)$$

Felírva a megoldóképletet és tovább rendezve:

$$p_{1,2} = \frac{2nx + c^2 \pm \sqrt{(2nx + c^2)^2 - 4(n + c^2)nx^2}}{2(n + c^2)} =$$

$$= \frac{nx + c^2 \cdot 0,5}{n + c^2} \pm c \sqrt{\frac{n}{(n + c^2)}} \sqrt{\frac{n \frac{x(1-x)}{n} + c^2 \frac{0,25}{n}}{(n + c^2)}} \quad (3.19)$$

Bevezetve a $\lambda = \frac{n}{n + c^2}$ (pozitív, 1-nél kisebb számmal való) helyettesítést:

$$p_{1,2} = [\lambda x + (1 - \lambda) \cdot 0,5] \pm z_{1 - \frac{\alpha}{2}} \cdot \sqrt{\frac{N - n}{N - 1}} \sqrt{\lambda} \sqrt{\left[\lambda \frac{x(1-x)}{n} + (1 - \lambda) \frac{0,5 \cdot (1 - 0,5)}{n} \right]} \quad (3.20)$$

Vegyük észre, hogy $n \rightarrow \infty$, és $\frac{n}{N} \rightarrow 0$ esetén $\lambda \rightarrow 1$ és $\frac{N - n}{N - 1} \rightarrow 1$, így becslésünk

határértékben az $x \pm z_{1 - \frac{\alpha}{2}} \cdot \sqrt{\frac{x(1-x)}{n}}$ alakot ölti (ez megegyezik az M1 becsléssel).

Annak ellenére, hogy egy meglepően szép konvex kombinációt tartalmaz a (3.20) becslőfüggvény, nehéz volna intuitív módon előállítani. A fejezet későbbi szakaszában bemutatott elemzésből azonban ki fog derülni, hogy az elméleti konstrukcióval összhangban kis minták esetén is gyakorlatilag torzítatlan (azaz a névlegesnek közel megfelelő megbízhatóságú) intervallumokat ad.

Visszatevés nélküli minta – egzakt hipergeometriai eloszlás (M3)

Egzakt becslést (a továbbiakban: M3) készíthetünk implicit megközelítéssel a hipergeometriai eloszlás adott sokasági hibaszám melletti gyakoriságfüggvényének m szerinti kumulálásával. Noha ez a kumulálás korábban – számításigényessége miatt – nem volt kivitelezhető, a számítástechnika folyamatos fejlődése jelenleg már lehetővé teszi az ilyen egzakt becslések alkalmazását. Az ellenőrzési szakma egyik „etalonnak” tekinthető szoftverének, az IDEÁ-nak⁵⁷ a mintavételi modulja is ezt az eljárást használja.

Mivel az így készült becslés egzakt becslés, így megbízhatósága „nem kérdés”. Az egzakt becslésekkel kapcsolatosan a fő kérdés az, hogy lehetséges-e ésszerű időn belül elvégezni a számításokat, illetve, hogy a közelítő becsléshez képest kapott többletpontosság megéri-e.

⁵⁷ Az IDEA szoftver neve egy rövidítésből származik (*Interactive Data Extraction and Analysis*), eredetileg a kanadai könyvvizsgálók részére kifejlesztett adatbázis-kezelő alkalmazásnak indult, de a nagy siker miatt később világszerte elterjedt. Magyarországon először a '90-es végén kezdte el alkalmazni az Állami Számvevőszék és a Kormányzati Ellenőrzési Hivatal, de nagy magáncégek, például a KPMG is használják.

Bayesi szemléletű becslések

A konfidencia intervallum becslés bayesi szemléletű definícióját, illetve a becslésre adott konstrukciós eljárást megvizsgálva látható, hogy a teljes valószínűségi mező intervallumokkal való lefedettségének mértéke általában nem egyenletes az egyes sokasági arányok mentén: ez amiatt van, hogy a priorban szereplő eloszlásnak megfelelően bizonyos sokasági arányok túltreprezentáltak a becslésben. Mindebből következik az is, hogy elvileg a bayesi szemléletű becselőfüggvények jóval pontosabb (szűkebb) intervallumokat eredményeznek; más kérdés, hogy ezen intervallumok ex-post megbízhatóságát hogyan befolyásolja, ha a prior jelentősen eltér a valóságtól.

Mivel az $F(m | M)$ feltételes eloszlás minden M -re olyan, hogy a hozzá tartozó feltételes sűrűségfüggvény n növekedésével egyre inkább egy pontra (és pedig $\frac{nM}{N}$ -re) koncentrálódik,

a priortól függetlenül minden intervallumbecslés $M = \frac{Nm}{n}$ értékre fog ráhúzódní, ha $n \rightarrow N$.

Kérdés azonban, hogy vajon ez a konvergencia milyen gyors, tehát hogy az ellenőrzésben használatos mintaméretek mellett érezteti-e a hatását. Fontos tudni, hogy a prior esetleges helytelen megválasztása milyen mértékben képes befolyásolni magát a becslést, és így az auditor által kialakított véleményt is.

Egy bayesi szemléletű intervallumbecslés elkészítéséhez két dolog ismerete szükséges: egyrészt ismernünk kell az $F(m | M)$ feltételes eloszlást, másrészt pedig az M prior eloszlását. Az ismert statisztikai összefüggések miatt $F(m | M)$ -et a hipergeometriai eloszlás értékeit felhasználva kaphatjuk meg, ám M eloszlásával kapcsolatosan többféle feltevéssel is élhetünk (ezek közül az első három, főleg az első kettő, inkább csak interpretációjában tér el):

(B1a.) Az első lehetséges feltevés, hogy *makro szemléletben* annak a valószínűsége (P_1), hogy egy adott egyed rendelkezik azzal a bizonyos jellegzetességgel, a korábbi ellenőrzési tapasztalatokból ismert, rendszersajátosságot tükröző, stabil paraméter. Egy ilyen – végtelen nagyságú – szuperpopulációt feltételezve a vizsgált sokaságra jellemző sokasági arány valószínűségi változó, amely (N, P_1) paraméterű binomiális eloszlást követ;

(B1b.) Másik lehetséges feltevés a vizsgált szervezetnél, vagy hasonló szervezeteknél lefolytatott korábbi ellenőrzések empirikus tapasztalatain alapul. Ezek alapján szintén megadható annak a valószínűsége (P_2), hogy egy adott egyed rendelkezik azzal a

bizonyos jellegzetességgel, így a sokasági arány ismét csak binomiális eloszlású, (N, P_2) paraméterekkel;

(B2.) Ha egy adott jellemző sokasági arányáról hosszú időre visszamenően rendelkezünk megfigyelésekkel, és úgy találjuk, ez a sokasági arány eloszlásában stabil, akkor ezt az empirikus adatsort is felhasználhatjuk a becslésünkhöz;

(B3.) Végül használhatunk nem informatív priorként egyenletes eloszlást.

Az első két feltevés mellett a következő posterior gyakoriságot kapjuk (ez értelemszerűen megegyezik a pontbecslésnél is használt (3.2) formulával):

$$f_m(M) = \frac{\binom{N}{M} P^M (1-P)^{N-M} \binom{M}{m} \binom{N-M}{n-m}}{\sum_{i \in [m; N-n+m]} \left\{ \binom{N}{i} P^i (1-P)^{N-i} \binom{i}{m} \binom{N-i}{n-m} \right\}}, \text{ ha } M \in [m; N-n+m], \text{ egyébként } 0.$$

Hasonlóan a pontbecsléshez, a posterior gyakoriság itt sem változik, ha M és m azonos módon változik, így fennáll a becslt alsó és felső intervallumokra fennáll az $M_a(m) = M_a(0) + m$ és $M_f(m) = M_f(0) + m$ összefüggés.

Az empirikus adatsor alapján minden sokasági hibaarányhoz hozzárendelhető egy $\text{Pr}_e(p)$ empirikus bekövetkezési valószínűség, aminek felhasználásával megadható a sokasági hibaszám prior eloszlása is. A posterior valószínűségek:

$$f_m(M) = \frac{\text{Pr}_e\left(\frac{M}{N}\right) \binom{M}{m} \binom{N-M}{n-m}}{\sum_{i \in [m; N-n+m]} \left\{ \text{Pr}_e\left(\frac{M}{N}\right) \binom{i}{m} \binom{N-i}{n-m} \right\}}. \quad (3.21)$$

Egyenletes eloszlású prior mellett a posterior gyakoriság a következő:

$$f_m(M) = \frac{\binom{M}{m} \binom{N-M}{n-m}}{\sum_{i \in [m; N-n+m]} \left\{ \binom{i}{m} \binom{N-i}{n-m} \right\}}, \quad (3.22)$$

ha $M \in [m; N-n+m]$, egyébként 0.

A könnyebb kezelhetőség kedvéért levezeték egy gyors közelítő módszert a B1 becsléssel (azaz binomiális priorral) kapható intervallumok meghatározására. Mivel fennáll az $f_m(M) = f_{m+x}(M+x)$ összefüggés, és $M-m=0$ melletti posterior eloszlása $(N-n, P)$

paraméterű binomiális eloszlás (lásd a (3.6) formulát), ezért M feltételes posterior várható értéke $(N-n)P + m$, feltételes posterior szórása pedig $\sqrt{(N-n)P(1-P)}$. A standard eljárás szerinti konfidencia-intervallum M -re:

$$(N-n)P + m \pm z_{1-\frac{\alpha}{2}} \sqrt{(N-n)P(1-P)} \quad (3.23)$$

Mivel minket nem M , hanem M/N , a sokasági hibaarány érdekel, ezt osztjuk N -nel, és azonnal alakítunk is rajta egy kicsit:

$$\frac{(N-n)}{N}P + \frac{n}{N} \frac{m}{n} \pm \frac{(N-n)}{N} z_{1-\frac{\alpha}{2}} \sqrt{\frac{P(1-P)}{(N-n)}} \quad (3.24)$$

Bevezetve a $\kappa = \frac{n}{N}$ jelölést a kiválasztási arányra a bayesi „gyorsbecslés” képlete:

$$\kappa \cdot \frac{m}{n} + (1-\kappa) \cdot \left(P \pm z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{P \cdot (1-P)}{N-n}} \right). \quad (3.25)$$

A (3.25) képlet a következő módon interpretálható: A sokaságból kiválasztott n elemű mintában $\frac{m}{n}$ a hibaarány. A mintavételkor „kihagyott” részben található hibák száma előzetes

feltevésünk szerint $(N-n, P)$ paraméterű binomiális eloszlást követ, erre vonatkozóan

$P \pm z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{P \cdot (1-P)}{N-n}}$ ad egy közelítő konfidencia-intervallumot. A becslést ennek a két

résznek a kiválasztási arány szerinti átlagolásával kapjuk. (Ez a becslés tehát csak azért közelítés, mivel a binomiális rész kvantiliseit közelítve határoztuk meg.)

Figyeljük meg, hogy a különböző kiválasztási arányok esetén hogyan viselkedik a (3.25) képlet. Ha nem veszünk mintát, akkor a becslés teljes mértékben a prior információn alapul,

hiszen $\kappa = \frac{n}{N} = 0$. Ahogyan nő a mintaméret, annál kisebb súlyt kap a becslésben a prior

információ, és annál nagyobbat a mintából leszűrhető. Abban a szélsőséges esetben, amikor 100%-os mintát veszünk ($\kappa = 1$), teljesen figyelmen kívül hagyjuk a priort.

Egy vegyes becslési módszer

A mintavételi statisztika keretében tárgyalt nagymintás becslési módszerből külső információ felhasználásával egy kis mintákra hatékonyan alkalmazható „hibrid” (MxB1) hozható létre.

Emlékeztetőül: az M1 nagymintás becslési eljárás során p mintabeli hibaarányhoz

hozzárendeltük a $p \pm z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{p \cdot (1-p)}{n}}$ intervallumot, ez azonban abban az esetben, ha a mintában nem találtunk hibát, egy nyilvánvalóan használhatatlan becslést ad. Ám ha előzetes tapasztalatokkal (P_0) rendelkezünk a hibaarányról, ezt a becslés során felhasználhatjuk a standard hiba meghatározásánál.

Az új becslőfüggvényünk:

$$p \pm z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{P_0 \cdot (1-P_0)}{n}}, \quad (3.26)$$

Ez a becslés bayesi intervallumbecslésként is felfogható azzal a feltételezéssel, hogy a sokasági hibaarány prior eloszlása a P_0 pontra koncentrálódik.

Nyilvánvaló, hogy ez a becslőfüggvény torzított, és a torzítás aszimptotikusan sem szűnik meg. Azonban kis minták esetén, ha P valóban P_0 közelében szóródik, ez a becslés jóval megbízhatóbb, mintha a standard hibát is a mintából határoznánk meg.

Az egylépcsős becslési eljárások értékelése

Noha a leírt becslési eljárások képletei önmagukban is „szépen mutatnak”, ismertetésüket nem ez indokolja. A kutatás célja az ellenőrzési gyakorlat szolgálata, jelen esetben egy elemzés elvégzése annak megítélésére, hogy az ellenőrzések során előforduló aránybecslési feladatokhoz mikor melyik módszert érdemes használni. Mivel az ellenőrzés célja a bizonyosság nyújtása, az intervallumbecslési (hipotézisvizsgálati) feladatoknál egyáltalán nem közömbös, hogy az alkalmazott módszer mennyire megbízható a valóságban. Ha az elvárt névleges szintnél valójában jóval magasabb a megbízhatósági szint, akkor ez azt jelenti, hogy felesleges munkát végeztünk, ha viszont alacsonyabb, akkor akaratunkon kívül nagyobb kockázatot vállaltunk. Ebben a szakaszban tehát értékelni fogom az eddig leírt becslési eljárások megbízhatóságát, nevezetesen a valós és a névleges megbízhatósági szintek közötti eltérést, amit az intervallumbecslés torzításának szokás nevezni.

Az értékelés módszere

A becslési eljárások értékelésekor *első lépésben* különböző sokaság- és mintaméretek mellett 15% alatti sokasági hibaarányokra Excel segítségével meghatároztam minden egyes sokasági

hibaarány esetén a megbízhatósági szinteket (a továbbiakban: parciális megbízhatósági szintek). A B1 és MxB1 eljárások esetén a sokaság- és mintaméret mellett paraméterként szerepelt a feltételezett hibaarány is, természetesen más-más interpretációval. A függelékben található ábra-gyűjteményen belül a 16. ábra mutatja be az egyik Excel tábla felépítését. (Természetesen a háttérben segédtáblák egész sora számította az itt megjelenő értékeket).

Sokaság mérete (N=)	Minta mérete (n=)	Hibaarány-prior (p=)
200	30	5%
200	50	5%
200	150	5%
600	30	5%
600	50	5%
600	150	5%
1500	50	5%
1500	200	5%

9. Táblázat: az elemzett forgatókönyvek

A becslőfüggvények megbízhatóságának elemzéséhez explicit módon meghatároztam, hogy a becsült intervallumok (m hipergeometriai eloszlását feltételezve) az esetek hány százalékában tartalmazzák M-et (illetve a valós sokasági hibaarányt). Az M3 módszert randomizálással tettem valóban torzítatlanná.⁵⁸ Ezzel párhuzamosan meghatározásra került egy várható intervallumhossz is minden becsléshez minden sokasági hibaarányra. A parciális megbízhatósági szinteket, illetve a becsült intervallum várható hosszát mutatja a grafikonok első csoportja (18. ábra – 33. ábra) a következő forgatókönyv-kombinációkra (az értékeléshez természetesen ennél többet kellett elemezni, de már ezek az ábrák is jól mutatják a tendenciákat).

A mintavételi statisztika becsléseinek torzítatlanság esetén minden sokasági arány mellett 95%-os parciális megbízhatóságot kell(ett volna) mutatniuk, ezzel szemben a bayesi szemléletű becslésektől, különösen a B1 becsléstől, ez nem várható el.

⁵⁸ Noha egzakt becslésről van szó, a torzítatlanság egyenlőségként (és nem nagyobb-egyenlőként) való értelmezése esetén diszkrét problémáknál általában az egzakt becslések sem torzítatlanok.

A randomizálást úgy vettem figyelembe az elemzés során, hogy minden sokasági hibaarányánál a tartalmazást kódoló 0-1 számok helyett a randomizált becsléshez adott *tartalmazási valószínűséggel* számoltam. Amennyiben nem randomizáltam volna, az M3 becslésnél pozitív torzítást kellett volna tapasztalni pár helyen (ugyanez a pozitív torzítás okozza például a többi grafikon „fűrészfogait” is).

Az értékelés *második szakaszára* éppen ezért volt szükség, tehát mivel a B1 becslés megbízhatóságának priortól való függése az első szakasz eredményeiből közvetlenül nem volt megállapítható. A második szakaszban tehát az első szakasz eredményeit a szuperpopulációs hibaarány 0-15% közötti értékeire kiszámított $\text{Pr}(M)$ súlyokkal átlagoltam. Egy ilyen esetet mutat a 17. ábra.

A bayesi szemléletű B2 becsléscsaládról nem tudtam univerzális érvénnyel nyilatkozni, de mivel ez a becslés átmenetet képez a „tisztá” B1 és B3 esetek közt, tulajdonságai is várhatóan valahol a kettő között találhatók: minél kevesebb információval rendelkezünk a jellemző előfordulási valószínűségéről – minél kevésbé összpontosul az empirikus eloszlás egy értékre – annál inkább a B3 jellemzői érvényesülnek, ami egyenletesebb parciális megbízhatóságot de hosszabb (pontatlanabb) intervallumokat jelent.

Az első szakasz eredményei

Az eredményekből kitűnt, hogy egyrészt az M2 becslés torzítása (a 95%-tól való eltérés) a többségében nem, vagy alig lépi túl a 1,5-2 százalékpontot, másrészt a torzítás hol pozitív, hol negatív irányú. Ezzel szemben az ellenőrzési szakirodalomban általában ajánlott M1 becslés akár 5-15 százalékpontnyi torzítást is tartalmaz. Az eredményeket tartalmazó grafikonokon a valós sokasági hibaarány növekedésével meglepő periodicitás figyelhető meg az M1 és M2 becsléseknél rögzített N és n mellett. Ezen kívül megfigyelhető, hogy a becslések megbízhatósági görbéjének alakja gyakorlatilag érzéketlen a sokaság nagyságának változásaira, amiből az a következtetés vonható le, hogy a megbízhatóságot elsősorban a mintaméret befolyásolja (a kiválasztási arány növekedése a növekedéssel arányos megbízhatósági szint növekedést okoz az M1 becslésnél⁵⁹, de a grafikon alakját érdemben sem az M1, sem az M2 becslésnél nem változtatja meg).

A mintavételi statisztika M3 egzakt módszer, nem tartalmaz semmilyen közelítő eljárást és ez esetben (a randomizálás miatt) torzítást sem, ezt az elemzés is alátámasztotta. Ha nincs randomizálás, a probléma diszkrét jellege miatt természetesen egyes sokaságoknál a megbízhatósági szint növekedését, pozitív torzítást tapasztalnánk.

⁵⁹ Az M1 becslésnél tapasztalt túlzott megbízhatósági szint növekedés például a statisztikai szakirodalomban ajánlott módon – a pontbecsléshez hozzáadott-kivont értékek (1-kiválasztási arány) tényezővel való szorzásával – korrigálható.

Az informatív priorral való B1 becslés a szuperpopulációs hibaarány környezetében pozitív, azon kívül negatív irányban torzított. A pozitív torzítási tartomány nagysága fordítottan arányos a sokaság méretével, hozzávetőlegesen 2,5 (N=1500) és 5 (N=300) százalékpont között mozog. A mintaméret (kiválasztási arány) változása a pozitív torzítási tartomány méretét kevésbé, inkább a (parciális) megbízhatósági görbe alakját befolyásolja: minél magasabb a kiválasztási arány, a becslés (parciális) megbízhatósága priortól függetlenül annál jobban közelít a konstans 95%-hoz. Noha ez a konvergencia csak a 100%-hoz közeli kiválasztási arányoknál kezd igazán érezhető lenni, egy bayesi szemléletű becslésnél nem is elvárás, hogy a parciális megbízhatóság egyenletesen közel legyen az elvárt megbízhatósági szinthez.

A nem informatív prior használata melletti B3 becslés a mintavételi statisztika M2 becsléséhez nagyon hasonló eredményeket hozott.

Az MxB1 vegyes becslés torzítása a hipotézisben lévő tévedéssel ellentétes irányban változik: túlbecsült hibaarány esetén növekszik a megbízhatósági szint, alulbecslés esetén pedig 95% alá csökken. Hüvelykujjszabályként megjegyezhető, hogy a hipotetikus hibaarány 1 százalékpontos csökkentése a megbízhatósági szintet 2-3 százalékponttal csökkenti.

Megvizsgáltam azt is, hogy az adott forgatókönyv mellett az egyes becselőfüggvények milyen várható pontossággal (intervallumhosszal) dolgoznak. Az informatív priort tartalmazó B1 becslést leszámítva az egyes becslések intervallumhossza a 0-15% közötti tartományon nagyságrendileg megegyezett, forgatókönyvtől függően 5 és 25 százalékpont között mozgott⁶⁰.

A második szakasz eredményei

A B1 becslés megbízhatósági szintje a sokasági hibaarány binomiális eloszlása mellett akkor maximális, ha a feltételezett szuperpopulációs hibaarány megegyezik annak valós értékével. Ebben az esetben a megbízhatósági szint pontosan 95%. A valós hibaaránytól távolodva csökken a becslés megbízhatósága is: minél nagyobb a kiválasztási arány, a csökkenés is annál lassabb ütemű. Ennek ellenére 200-300 elemű minták esetén is nagyon erős a prior jóságától való függés.

⁶⁰ Elsőre meglepő lehet, hogy az elvileg fix pontosságú MxB1 becslésnek miért nem konstans a várható intervallumhossza. A jelenség magyarázata, hogy az intervallumoknak csak a nemnegatív szakaszát vesszük figyelembe.

Itt kell szót ejtenünk arról, hogy ezt az eredményt hogyan is kell pontosan értelmezni. A bayesi becslés valójában nem csinál mást, mint a sokaságról szerzett többletinformációt a pontosság növelésére fordítja. Ez az átváltás mindaddig jól működik, amíg a becslésbe bevitt többletinformáció (megközelítőleg) helyes. Ha azonban az információ pontatlan, súlyos árat fizetünk: drasztikusan csökkeni fog becslésünk megbízhatósága (persze lehet olyan szerencsénk, hogy a konkrét sokasági realizáció mégis nekünk kedvez, de erre nem érdemes alapozni).

Következtetések

Az eddig leírtakból levonható fontos következtetések:

- A legfontosabb az M1 becselőfüggvény alkalmazhatóságáról levonható következtetés: a becslés 5-10%-os várható hibaarány mellett az alkalmazásához legalább 150-200 elemű minta szükséges. Jól funkcionál az alapozó statisztika könyvekben alkalmazási feltételként javasolt $\min\{np, n(1-p)\} \geq 10$ hüvelykujjszabály, kiegészítve azzal, hogy a végesség miatt a szórásnál indokolt az $\frac{N-n}{N-1} \approx 1 - \frac{n}{N}$ korrekciós tényező használata.
- Az előbbi egyenes következménye, hogy a képlet „visszafelé történő alkalmazásával” számított szükséges mintaméretek sem megfelelőek, csak a nagyságrendek gyors kiszámítására alkalmasak. Ha tehát statisztikailag is megalapozott következtetéseket akarunk levonni a mintából, akkor ne használjuk az ellenőrzési szakirodalomban általánosan elterjedt $n = \frac{z^2 p(1-p)}{d^2}$ képletet a mintaméret meghatározására, ehelyett használjuk inkább az IDEA M3 eljárásen alapuló beépített mintavételi modulját.
- Kérdéses, hogy 10 százalékpont feletti becsült intervallumhossz mellett van-e egyáltalán értelme az intervallumbecslésnek, ha a tolerálható hibaarány 5-10% körül mozog.
- Amennyiben többéves, részletes idősoraink vannak egy adott hiba, szabálytalanság előfordulási gyakoriságára vonatkozóan, és ennek a hibaarányának az eloszlása időtől és/vagy vizsgált szervezettől függetlenül nagyfokú stabilitást mutat, érdemes megfontolni a bayesi intervallumbecslés használatát, ugyanis ez a becslés az elvárt megbízhatóság teljesítése mellett nagy pontosságú (rövidebb) intervallumokat szolgáltat.

- A bayesi megközelítés alkalmazása csak akkor elfogadható, ha valamilyen módszerrel biztosítható, hogy az alkalmazott prior közelítőleg helyes lesz. Ha ez nem áll fenn, akkor a bayesi becslések használata nagy, és nem számszerűsíthető kockázatot jelent.

Intervallumbecslés, hipotézisvizsgálat (szekvenciális mintavétel)

A szekvenciális mintavételi eljárás első részletes leírását Wald Ábrahám adta (Wald [1945]), az eljárás gyökereit maga a szerző is a kétlépcsős, illetve az általánosabb többlépcsős mintavételi eljárásokban látta. Az egylépcsős, kétlépcsős, többlépcsős és szekvenciális eljárások közötti logikai kapcsolat az alábbi módon foglалható össze egy tipikus ellenőrzési szituáció, a sokasági arányra vonatkozó tesztelés (acceptance sampling in attributes) esetén:

- Egylépcsős mintavétel esetén kiválasztunk egy n elemű mintát. A mintaelemek megvizsgálását követően összevetjük a mintában található minősített elemek (a továbbiakban az egyszerűség kedvéért hibák) számát (m) az előre definiált – a döntési szabályunkban szereplő – értékkel (a), és a két érték relációjától függően döntünk ($a \geq m$ esetén pozitív, $a < m$ esetén negatív döntést hozunk). A későbbiekkel való könnyebb összehasonlíthatóság kedvéért formálisan definiáljuk az $r = a + 1$ értéket, és írjuk át a negatív döntés feltételét $r \leq m$ alakba.
- Két-, illetve többlépcsős (k lépcsős) mintavétel esetén minden lépcsőnél meghatározunk egy n_i elemű mintát. Az utolsó, $i = k$. lépcsőt kivéve minden esetben $r_i > a_i + 1$. A döntést a korábbival analóg módon hozzuk: ha m , a hibák kumulált száma az i . minta után teljesíti az $a_i \geq m$ vagy az $r_i \leq m$ relációk egyikét, pozitív (acceptance), illetve negatív (rejection) döntést hozunk, és leállunk. Ha azonban $a_i < m < r_i$ az i . lépcső után, akkor a következő, n_{i+1} db tételen is lefolytatjuk az ellenőrzést. Mivel az utolsó lépcsőnél $r_k = a_k + 1$, ezért az eljárás biztosan befejeződik.
- A szekvenciális mintavétel olyan speciális többlépcsős mintavételi eljárás, ahol minden egyes lépcsőhöz egy egyelemű minta tartozik. Alapértelmezésben (és az optimális tesztnél) $r_i > a_i + 1$ minden i -re, de természetesen megadható felső korlát a lehetséges lépcsők számára. A pozitív és a negatív döntés szabálya itt is a szokásos ($a_i \geq m$ és $r_i \leq m$).

Az egylépcsős mintavételi eljárással szemben a két- vagy töblépcsős és a szekvenciális mintavételi eljárásnál nem lehet előre megmondani, mekkora lesz az ellenőrzendő minta (kevésbé tervezhető a munkaráfördítés), ezzel szemben bizonyítható, hogy várható értékben ugyanazt a bizonyosságot jóval kisebb mintamérettel is el lehet érni. Vegyük észre továbbá, hogy az összes többi eljárás felírható egy megfelelően paraméterezett szekvenciális eljárásként (tekintsük pl. az egylépcsős eljárást: a minta utolsó, n . tételét leszámítva minden korábbi tételnél $a_i = -1$ és $r_i = n + 1$, az n . tételre pedig $a_n = a$ és $r_n = a_n + 1$)

Mind a statisztikai, mind a köznapis intuición oldaláról könnyen látható a töblépcsős eljárások előnye az egylépcsős eljárással szemben. Amennyiben dönteni kell valamiről, az extrém rossz és extrém jó változatokat könnyen és gyorsan ki lehet szűrni, a probléma mindig a határesetekkel van, ahol azután a következő körben további vizsgálatokat fogunk lefolytatni. Statisztikailag megfogalmazva ugyanezt: a „jól viselkedő” próbák ereje a tesztelt hipotézistől távolodva általában nő. Ha nem mindenhol növeljük meg a mintaméretet, hanem csak a nullhipotézis közelében lévő eseteket vizsgáljuk meg alaposabban (értsd: nagyobb mintát véve), akkor ez által rengeteg felesleges munka takarítható meg, és a próba ereje is mindenhol el fogja érni az elvárhatót.

Mivel az egyszerű egylépcsős eljárással szemben rengeteg töblépcsős eljárás készíthető, felmerül a kérdés, melyik közülük az „optimális”. Ha az optimalitás kritériuma az ellenőrzött tételek számának várható értéke (tehát az a „legjobb” módszer, ahol várhatóan a legkevesebb tételt kell megnézni), akkor a szekvenciális mintavételi eljárás lesz az optimális a következő döntési szabállyal:

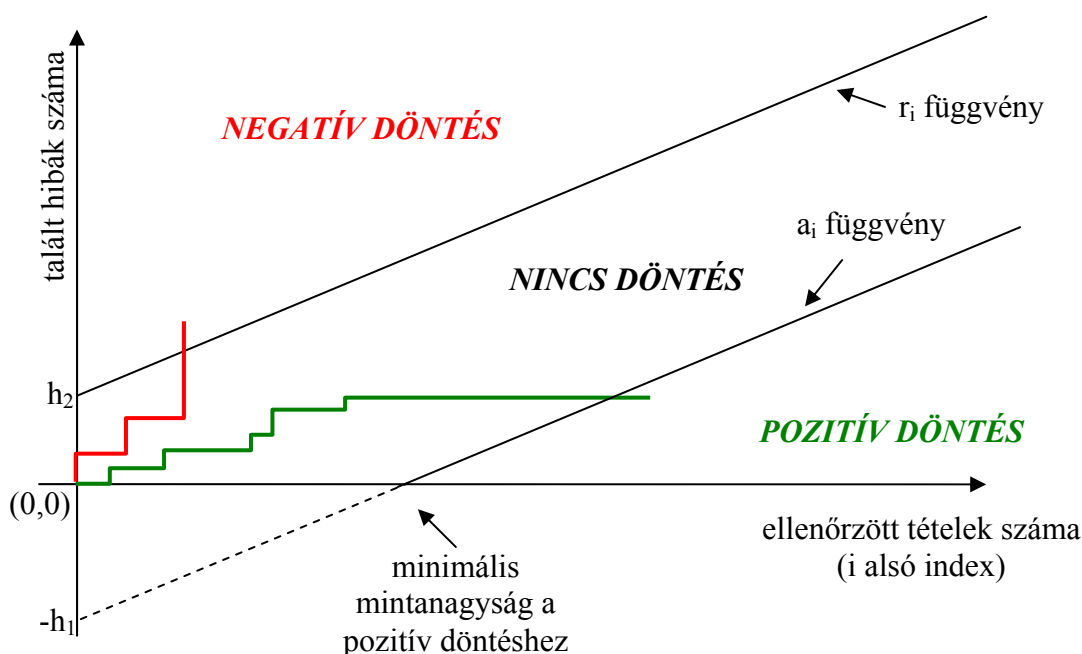
$$a_i = -h_1 + s \cdot i \text{ és } r_i = h_2 + s \cdot i, \text{ ahol} \quad (3.27)$$

$$h_1 = \frac{\ln \frac{1-\alpha}{\beta}}{\ln \frac{p_1(1-p_0)}{p_0(1-p_1)}}, \quad h_2 = \frac{\ln \frac{1-\beta}{\alpha}}{\ln \frac{p_1(1-p_0)}{p_0(1-p_1)}} \text{ és } s = \frac{\ln \frac{1-p_0}{1-p_1}}{\ln \frac{p_1(1-p_0)}{p_0(1-p_1)}}. \quad (3.28)$$

Az ilyen paraméterekkel felírt teszt paramétereinek jelentése: legfeljebb α a valószínűsége annak, hogy egy p_0 -nál nem nagyobb hibaarányú sokaságra negatív véleményt adunk, illetve legfeljebb β a valószínűsége annak, hogy egy p_1 -nél nem kisebb hibaarányú sokaságra pozitív véleményt adunk ($p_0 < p_1$).

Ezt a (3.27) képletben leírt döntési szabályt grafikusán ábrázolva megkapjuk a szekvenciális teszt „tipikus” ábráját. A 6. ábra alapján könnyen elképzelhető, hogyan is zajlik a teszt: a (0;0)

pontból indulva, ha az adott mintaméretnél a hibák száma eléri, vagy meghaladja a felső vonalat, negatív döntést hozunk (piros vonallal jelzett példa), ha eléri, vagy aláégy az alsó vonalnak, akkor pedig pozitív döntést (zöld vonallal jelzett példa). Mindaddig, amíg a két vonal között vagyunk, folytatjuk az eljárást, és megnézzünk még egy tételt. Figyeljük meg az ábrán, hogy a (3.27)-(3.28) képletekkel felírt (várható tételszám szempontjából) optimális döntési szabály – melyet az a_i és r_i függvények együttese reprezentál – elvileg lehetővé teszi, hogy a mintavétel a végtelenségig folytatódjon, aminek már az elvi lehetősége is kizárja az ellenőrzési alkalmazást.



6. ábra: A szekvenciális teszt lehetséges forgatókönyvei

Noha Wald bizonyította, hogy a szükséges mintaelemszám várható értéke szempontjából az ilyen szekvenciális tesztelés az optimális, többlépcsős eljárások alkalmazása nélkül is „van tartalék” az egylépcsős eljárásban (mintaelemszám-optimalizálási szempontból). Ha például valamelyik mintaelem ellenőrzését követően azt tapasztaljuk, hogy a kumulált hibák száma meghaladta a tolerálhatót ($m = r = a + 1$), akkor leállhatunk, hiszen kevesebb hiba már úgyse lesz. Hasonlóképpen, ha $n - a$ tétel átvizsgálását követően még nem találtunk hibás tételt, leállhatunk, hiszen a további „a” db. tételben legfeljebb a, azaz elfogadható mennyiségű hibát tudunk találni.

Az optimalizálást egy ellenőrzési alkalmazásnál nem feltétlenül szabad csupán a várható tételszámra elvégezni. Korábban már említésre került, hogy a Wald által definiált „optimális”

teszt eleve kizárja magát a könyvvizsgálati alkalmazásokból csak az által, hogy elméletileg akár a végtelenségig is folyhat a tesztelés. A másik probléma a Wald-féle optimális tesztrel, hogy a működéséről nehezen adható a statisztikát csak eszköznek tekintő alkalmazóknak intuitív magyarázat: noha az 50-es évektől kezdve évtizedenként megjelent 1-2 cikk, ami a szekvenciális módszer bevezetését javasolta a könyvvizsgálatban, a gyakorlatban továbbra sem használnak mást, szinte csak egylépcsős mintavételt.⁶¹ Az egylépcsős mintavétel előnye a többlépcsőssel szemben a könnyebb interpretálhatóság, tehát ha a gyakorlatban is használt, könnyen alkalmazható szekvenciális módszert szeretnénk, akkor annak mindenképpen az egylépcsős mintavételből kell kiindulnia; ez azonban egyben azt is jelenti, hogy kompromisszumot kell kötnünk a várható tételszámmal mért optimalitás kárára.

Az önkormányzatok átfogó ellenőrzéséhez kidolgozott mintavételi módszer

Az Állami Számvevőszéknél részt vettem abban a munkacsoportban, amely a következő 4 évben (2007-2010) a kiemelt, nagy költségvetésű önkormányzatoknál lefolytatandó átfogó ellenőrzések ellenőrzési programjának átdolgozását végezte. A csoport szakmai szempontok alapján meghatározta az ellenőrzendő kontroll eljárásokat, és megjelölte a kontrolloktól elvárt megbízhatóságot (a tolerálható hibaarány ellentettjét). A feladat egy olyan mintavételi-következtetéseméleti módszer kidolgozása volt, amely

- Ésszerű bizonyosságot szolgáltat arról, hogy a kérdéses kontroll valóban működik;
- Minimalizálja a kontrollokat megfelelően működtető ellenőrzöttek kockázatát, azaz a lehetőségekhez mérten nem „statáriális”;
- Minimalizálja az ellenőrzést végzők felesleges munkaterhelését.

Amennyiben csak az első két szempontot vennénk figyelembe, a megoldás magától értetődő volna: nagy mintát kell venni. A nagy minta ugyanis csökkenti mindkét hiba kockázatát. Ez azonban – figyelembe véve az egyéb ellenőrzési célokat is – túl sok (és főleg felesleges) munkaráfordítást jelentene.

A célok összehangolását nagyban segítette a helyszíni ellenőrzést végző kollégák tapasztalata, mely szerint az ellenőrzöttek vagy nagyon jók, vagy nagyon rosszak, és kevés az olyan

⁶¹ Ellenkező esetben a szekvenciális módszer megjelenne a standard oktatási anyagokban is, amiből a könyvvizsgálók 99,99%-a az ismereteit szerzi.

szervezet, amelyik éppen a határon mozog. Mindez arra utalt, hogy valamilyen többlépcsős, vagy szekvenciális jellegű eljárást kell kidolgozni.

A könnyebb interpretálhatóság miatt a negatív megállási pont (tehát az r_i függvény) úgy lett kialakítva, hogy az eggyel nagyobb legyen, mint a sima egylépcsős mintában az eredeti kiindulási feltételek mellett éppen tolerálható hibák száma. Ilyen módon a negatív megállási pont interpretálása egyszerű: ha a teljes mintát ellenőriztük is volna, akkor sem juthatnánk más döntésre (Wald negatív megállási pontjainak interpretálása ennél jóval nehezebb).

A pozitív megállási pontok száma megegyezik az egylépcsős mintavétel esetén tolerálható hibák számával plusz eggyel (0-tól az utolsó tolerálható hibáig). A megállási pontok meghatározásához fel kell írni annak az eseménynek a valószínűségét, hogy a sokaság tolerálható hibaaránya mellett ezzel a mintavételi eljárással pozitív döntést fogunk hozni: definíció szerint ennek a valószínűségnek kell megegyeznie az ellenőrzés feltárási kockázatának előre rögzített valószínűségével (a mintavételes tesztelés megbízhatósági szintjének komplementerével).

A keresett valószínűség kiszámítási módjának bemutatásánál az alábbi jelöléseket használjuk:

- A tolerálható hibák száma a mintában: k ;
- A pozitív megállási pontok indexei: $0, 1, \dots, k$, tehát összesen $k+1$ darab pozitív megállási pontunk van;
- A pozitív megállási pontok között lévő részminták elemszáma: $n_0, n_1, \dots, n_k$, ezek összege adja meg a maximálisan megvizsgálandó tételek számát, ennek megfelelően a szekvenciális mintavétel lehetséges pozitív megállási pontjai: $h_0 = n_0, h_1 = n_0 + n_1, \dots, h_k = n_0 + n_1 + \dots + n_k$;
- A pozitív megállási pontok között talált hibás tételek száma: $m_0, m_1, \dots, m_k$, ezek kumulálásával kapjuk az adott megállási pontig felderített hibás tételek számát.

Jelentse P_i annak az eseménynek a valószínűségét, hogy az i . megállási pontig összesen i hibát találtunk ($\sum_{j=0}^i m_j = i$), és a $0, 1, \dots, i-1$ indexű megállási pontok mindegyikénél a hibák

kumulált száma nagyobb volt a megállási pont indexénél ($\sum_{j=0}^s m_j > s, \forall s \in [0; i-1]$). Ez az esemény egyenértékű azzal, hogy a tesztelés során pontosan az i . megállási pontnál fogjuk

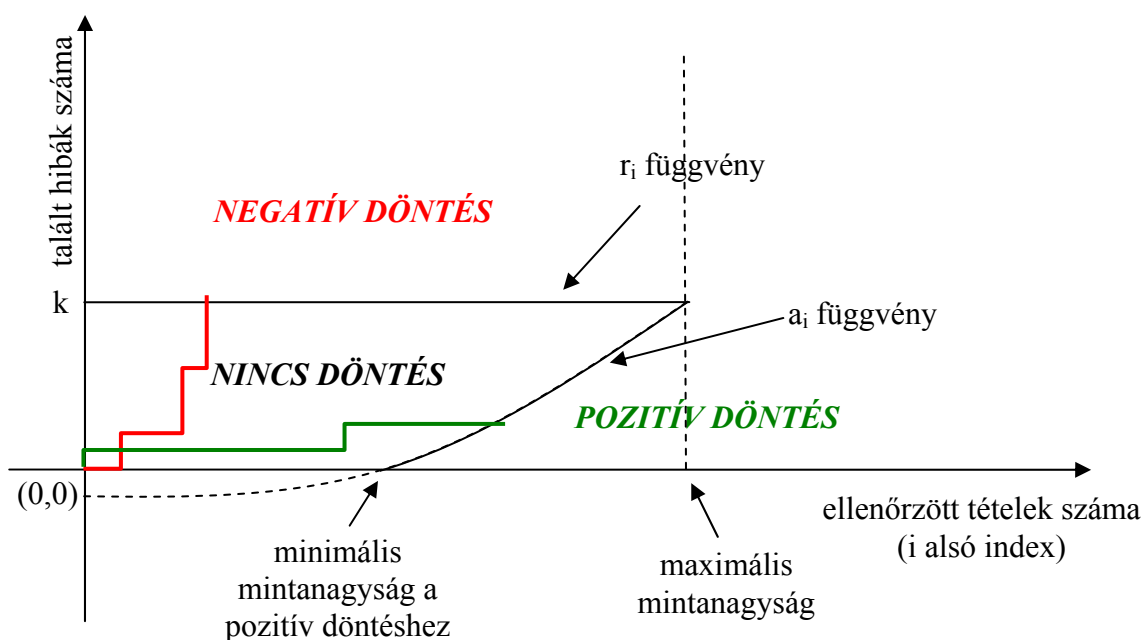
abbahagyni a tesztelést pozitív véleménnyel. A feltárási kockázat keresett valószínűsége tehát

általában: $\sum_{i=0}^k P_i$.

A P_i valószínűségek meghatározása annál bonyolultabb, minél nagyobb az i index. A gyakorlati kivitelezés során erre a célra érdemes egy Excel makrót készíteni, ami végignézi legfeljebb i db hiba összesen $\binom{2i+2}{i+1}$ db lehetséges elosztását a megállási pontok között, és ezekből kiértékeli azokat, amelyek a P_i valószínűség meghatározásához relevánsak.

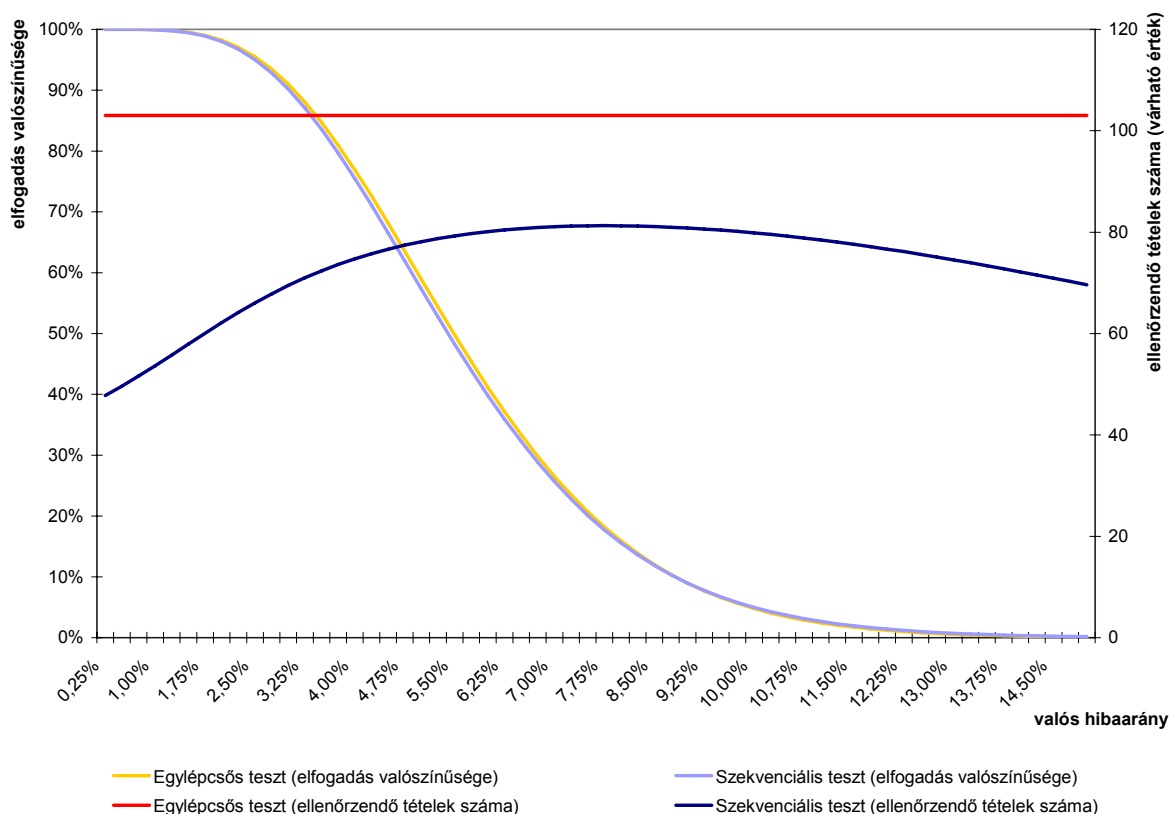
Visszatérve a pozitív megállási pontok meghatározásához, a feladat úgy meghatározni a h_i pontokat, hogy a $\sum_{i=0}^k P_i$ valószínűség a tolerálható sokasági hibaarány mellett egyenlő legyen

az előre definiált bizonyossági szint komplementerével, a szóban forgó alkalmazás esetén például 5%-kal (95% bizonyossággal akarunk nyilatkozni a tesztelt kontroll megbízhatóságáról). Ez valójában azt jelenti, hogy valahogyan szét kell osztanunk a kockázatot a megállási pontok között. Az Állami Számvevőszéknél kifejlesztett módszer ezt egyenlően osztotta szét (tehát minden pozitív megállás esetén ugyanakkora volt a kockázata az elsőfajú hiba bekövetkezésének). Az egyenlő szétosztás nem feltétlenül minimalizálja a mintaméretet, de az optimum numerikus meghatározása aránytalanul hosszú ideig tartott az esetlegesen elért mintaméret-csökkenéshez viszonyítva.



7. ábra: A módosított szekvenciális teszt lehetséges forgatókönyvei

A 6. ábra és a 7. ábra összehasonlításával válik igazán érthetővé, hogy miben más ez a módszer, mint a standard szekvenciális. Az első szembevetendő különbség, hogy a negatív döntés határa itt vízszintes, és az egylépcsős minta tolerálható hibájánál (k) húzódik (pont a határon értelemszerűen ez esetben még nincs negatív döntés). A másik különbség a pozitív döntés határának alakja: amennyiben a megállási pontok között egyenlően osztjuk szét a kockázatot, a tapasztalatok szerint határ nem lineáris, hanem egy konvex görbe. Ha a szétosztásnál az első egy-két pozitív megállási pontra (0 hiba, 1 hiba, esetleg 2 hiba) az egyenlő szétosztáshoz képest több kockázatot hagynánk, a nagyobb kockázatvállalás értelemszerűen alacsonyabb értéket fog adni ezekre a megállási pontokra, ami egy olyan sokaság esetén, ahol sok a jó és rossz, de kevés a közepes egyed, elvben még tovább növelheti a módszer hatékonyságát. Az utolsó különbség a két ábra között, hogy itt megadható a minta maximális tételszáma – ez az érték használható egyben egy hasonló egylépcsős teszt mintaméreteként is.



8. ábra: Az egylépcsős és a szekvenciális teszt összehasonlítása (a mintában tolerálható hibák száma: 5, tesztelendő hipotézis: 10%, szignifikancia szint: 5%)

A 8. ábra egy konkrét egylépcsős-szekvenciális teszt pár esetén megmutatja, hogy mennyi az esélye a hipotézis elfogadásának, illetve hogy mekkora az ellenőrzendő tételek számának

várható értéke pozitív vélemény esetén (negatív vélemény esetén a módszer csupán az egylépcsős módszer „természetes tartalékait” használja fel). Az ábráról leolvasható, hogy ugyan a szekvenciális teszt ereje pár százalékpontnyit elmarad az egylépcsős tesztétől a legtöbb ellenhipotézis esetén, azonban a szükséges mintaméretben várható megtakarítás a legrosszabb esetben is legalább 20%. A valós helyzetben ezt még módosítja az a megfigyelés, hogy a vizsgált egységek vagy nagyon jók, vagy nagyon rosszak, ennek megfelelően – amint az a következő szakasz adatfeldolgozásából is kitűnik majd – a valóságban tapasztalt megtakarítás ezt az értéket jelentősen meghaladja.

A módszer alkalmazási tapasztalatai a 2007. év első félévében

Az önkormányzatok gazdálkodási rendszerének 2007. évi átfogó ellenőrzésénél az új ellenőrzési programmal együtt bevezetésre került a szekvenciális (azaz megállási pontokat alkalmazó) mintavételi eljárás alkalmazása. Az eljárás újszerűsége miatt a bevezetés jelentős előkészületeket igényelt. Ennek keretében:

- kidolgozásra került egy standard Excel munkalap (tesztlap), amely teljes mértékben dokumentálja és reprodukálhatóvá teszi a mintavétel folyamatát, valamint támogatja az ellenőrzést végzőt a tesztelés lefolytatásában;
- a folyamatot két szakaszra bontottuk: minta kiválasztás és minta kiértékelés. Tekintettel arra, hogy a minta kiválasztása és a megállási pontok meghatározása speciális szakismeretet igényel, célszerűnek láttuk a folyamat ezen szakaszát központosítani, így csökkentve mind az oktatási igényt, mind az esetleges téves alkalmazások kockázatát. A mintavétel IDEA segítségével történő lebonyolítására és a megállási pontok kiszámolására az Önkormányzatok Gazdálkodási Rendszerét Ellenőrző Osztályon 2 főt „képeztünk ki”;
- a helyszíni ellenőrzést végzőket az év eleji vizsgálatindító oktatáson tájékoztattuk arról, hogy pontosan mit kell tenniük a helyszínen: elkérik a szükséges fájlokat az ellenőrzött önkormányzattól, azokat felküldik a kijelölt közreműködőknek, onnét pedig már a kiválasztott mintát tartalmazó tesztlapot kapják vissza, amin lefolytathatják az ellenőrzést.

Annak érdekében, hogy a felesleges adminisztrációt minimalizáljuk, azokban az esetekben, ahol a sokaság mérete nem haladta meg az 50-et, eltekintettünk a mintavételtől, és tételes ellenőrzést írtunk elő.

Az oktatáson különösen kihangsúlyoztuk, hogy mivel ugyanazon a bizonylaton több kontrollt is ellenőriznek, ezért ha az egyik kontroll tekintetében elérik a megállási pontot, akkor annak a kontrollnak a működését a továbbiakban már **nem szabad** ellenőrizni, abban az oszlopban **meg kell állni**. Ennek az elfogadása – érthető okokból – különösen nagy problémát okozott sok számvevőnek, ennek következtében az elején sok, helytelenül kitöltött munkalap érkezett be a minőségi felülvizsgálatra.

További problémát jelentett az is, hogy az ellenőrök többsége „nem akarta elhinni”, hogy a megállási pontok mellett is ugyanazt a bizonyosságot kapják, ezért sokan figyelmen kívül hagyták a megállási pontot, és a teljes mintát megvizsgálták. Remélhetőleg erről a (kicsit túlzott) professzionális szkepticizmusról előbb-utóbb leszoknak a kollégák, látva, hogy ugyanazt az eredményt kapják kevesebb tétel ellenőrzése után is.

A hibás könyvelés miatt gyakori gond, hogy a kapott állományokból vett minta tartalmaz irreleváns tételeket. Mivel ezek a tranzakciók nem vehetők figyelembe a vizsgált kontroll szempontjából, de a tévességük csak a helyszíni ellenőrzés során derül ki, ezért a mintavételt végzők már előre gondoskodnak „pót-tételekről” (20%). Ha a nem oda tartozó bizonylatok száma túl magas, az a könyvelési hiba rendszerszerűségére utal, ami innentől az egész ellenőrzést más megvilágításba helyezi.

A korábbi gyakorlattal ellentétben, amikor is tipikusan előre megadott időszakot, vagy hasonló, az irattárból könnyen kiválogatható mintát néztek a kollégák, az éven belüli elszórt tranzakciók kiválogatása az ellenőrzött szervezetnél jelentős többletmunkát okoz. Nem szabad azonban megfeledkezni arról sem, hogy ez a fajta „összevisszaság” sokkal jobban reprezentálja az egész évet, mint ugyanannyi tétel egy előre megadott adott időszakból történő kiválasztása.

A megállási pontok alkalmazásától azt reméltük, hogy csökkenni fog a bizonylatok ellenőrzéséhez szükséges idő. A beérkezett ellenőrzések tesztlapjainak statisztikai feldolgozása során figyelmen kívül hagytuk azokat az eseteket, ahol a kontroll 10, vagy annál kevesebb tranzakciót érintett. Ezt figyelembe véve a feldolgozott ellenőrzéseknél összesen 8623 tétel került kiválasztásra, amit a hagyományos eljárás esetén valószínűleg mind meg kellett volna vizsgálni. Mivel az 50 tétel alatti sokaságoknál csak negatív megállási pont lehetséges a tételes ellenőrzés miatt, ezért az elemzés során külön is megvizsgáltuk az 50 tétel feletti és alatti sokaságokat (ezek az összes vizsgált sokaság körülbelül felét-felét tették ki, viszont a belőlük vett minták tételszámát tekintve már 70-30 az arány, ahol a kisebb rész jelenti értelemszerűen az 50 tétel alatti sokaságokat).

A 10 tranzakciónál nagyobb sokaságok esetén a vizsgált hat kontrollnál a következő összesített eredményeket kaptuk:

Mintába került tranzakciók száma összesen		A mintából leellenőrzött tranzakciók száma összesen az egyes kontrollok esetén					
		Kifizetés jogosultságának szakmai igazolása	Kifizetés összecszerülésének szakmai igazolása	Szerződés, megbízás, megrendelés szakmai teljesítésének igazolása	Gazdálkodásra vonatkozó szabályok betartásának ellenőrzése	Szakmai telj.ig. megtörténtének ellenőrzése	Érvényesítés megtörténtének ellenőrzése
50 tétel és alatta	2599	1453	1448	1440	1449	1462	1612
		55,91%	55,71%	55,41%	55,75%	56,25%	62,02%
50 tétel felett	6024	2050	2068	1961	1832	1905	1870
		34,03%	34,33%	32,55%	30,41%	31,62%	31,04%
Összesen	8623	3503	3516	3401	3281	3367	3482
		40,62%	40,77%	39,44%	38,05%	39,05%	40,38%

10. Táblázat: A módosított szekvenciális módszer alkalmazásának tapasztalatai

Látható, hogy a módszer alkalmazása különösen ott célravezető, ahol sok az ellenőrzendő bizonylat, ezekben az esetekben a megtakarítás mértéke elérheti a 65-70%-ot is. Az önkormányzatok vizsgált összetételű vegyes populációján az átlagos megtakarítás körülbelül 60%.

Igazolást nyert az a hipotézis is, amely szerint az önkormányzatok jelentős része vagy jó, vagy rossz, és kevés a „közepes”. Ez abból látszik, hogy az önkormányzatok 60%-ánál a tranzakciók kevesebb, mint 40%-a elégséges volt a vélemény kialakításához, ami arra utal, hogy ezek az önkormányzatok elég egyértelműen besorolhatók a jó/rossz kategóriába („közepes” önkormányzatok esetében nem jutnánk el ilyen hamar a megállási ponthoz).

A módszer ellenőrzésben való alkalmazása azért ígéretes, mivel az ellenőrzöttek szélsőséges jellege miatt várhatóan elég hamar meg lehet állni vagy pozitív, vagy negatív döntéssel.

IV. rész: Értékösszeg-becslésre visszavezethető ellenőrzési feladatok

A könyvvizsgálatban a leginkább tipikusnak nevezhető feladat az, hogy egy pénzügyi kimutatásról el kell dönteni, tartalmaz-e lényeges hibát. Egy hiba lényegessége (materiality) a korábban adott értelmezés szerint abból fakad, hogy miatta már érdemben módosulnak a pénzügyi kimutatás alapján hozott döntések. A jelen disszertáció, és a legfontosabb gyakorlati alkalmazások szempontjából a lényeges hiba mindig a főösszeg (vagy fontosabb részösszegek) kimutatott és valós értékének⁶² egy tolerálható mértéket meghaladó eltérését jelenti⁶³. Például: lényeges a hiba, ha az eltérés meghaladja az 1 millió forintot, vagy lényeges a hiba, ha az eltérés meghaladja a főösszeg 2%-át. Ezt az összeget (vagy százalékot) más néven lényegességi küszöbnek is hívják. Mivel semmiféle megszorítást nem jelent, a továbbiakban a lényegességi küszöb mindig a főösszegre vonatkozik, és a főösszeg százalékában (tehát nem abszolút összegben) van megadva.

Ez a definíció a következő példával érzékeltethető legkönnyebben. Tekintsük a számviteli törvényben leírt „A” típusú mérleg egy egyszerűsített formáját:

Eszközök (aktívák)	Források (passzívák)
A. Befektetett eszközök	D. Saját tőke
I. Immateriális javak	I. Jegyzett tőke
II. Tárgyi eszközök	...
III. Befektetett pénzügyi eszközök	VII. Mérleg szerinti eredmény
B. Forgóeszközök	E. Céltartalékok
I. Készletek	F. Kötelezettségek
II. Követelések	I. Hátrasorolt kötelezettségek
III. Értékpapírok	II. Hosszú lejáratú kötelezettségek
IV. Pénzeszközök	III. Rövid lejáratú kötelezettségek
C. Aktív időbeli elhatárolások	G. Passzív időbeli elhatárolások
Eszközök összesen	Források összesen

11. Táblázat: Az „A” típusú mérleg felépítése

⁶² A „valós érték” kifejezést nem a számviteli törvényben (2000. évi C. tv. 3.§. (9) 12.) definiált értelemben használom. A továbbiakban „valós érték” alatt az auditor által végzett *teljes körű* ellenőrzés után kapott „helyes” értéket kell érteni. Természetesen *ez az érték hipotetikus*, hiszen teljes körű ellenőrzésre nem kerül sor.

⁶³ A lényegesség, mint arról már korábban szó volt, az itt leírtnál jóval összetettebb fogalom, azonban a fent leírtan kívüli dimenzióit alapvetően nem a statisztika eszközeivel szokás megragadni.

Tegyük fel, hogy a mérlegben szereplő eszközök, mint kimutatás megbízhatóságáról kell véleményt nyilvánítani. Tegyük fel továbbá a példa kedvéért, hogy lényegesnek tekinthető az Eszközök összértékének 2 százalékos, a Befektetett eszközök értékének 1 százalékos, a Forgóeszközök értékének 5 százalékos, az Aktív időbeli elhatárolások értékének 2 százalékos, és végül a Készletek és a Követelések együttes értékének 10 százalékos eltérése. Az egyes lényegességi küszöböket külön-külön kell vizsgálni. Könnyen látható, hogy a Befektetett eszközök, az Aktív időbeli elhatárolások, illetve a Készletek és a Követelések együttes értékének vizsgálata „tiszta eset”, hiszen nem tartalmaznak további, lényegességi küszöböket tartalmazó alábontásokat. Ezzel szemben a Forgóeszközök, és az Eszközök összesen értékelése ebben az értelemben többlépcsős folyamat. Kezdetben csak a tiszta esetekkel foglalkozunk, és csak később térünk ki röviden az összetett esetek kezelési módjára.

A célom a továbbiakban azon klasszikus (Neyman-Pearson elvet követő) statisztikai eszközök ismertetése, melyeket az elmúlt 50 évben fejlesztettek ki és jelenleg is használnak a pénzügyi beszámolók megbízhatóságának megítéléséhez. Noha bayesi eszközök is léteznek erre a feladatra, alkalmazásuk koránt sem annyira elterjedt, mint az aránybecslésnél, így azokat nem ismertetem.

A problémát a statisztika nyelvére lefordítva a következő feladattal állunk szemben:

1. minta alapján becsülni kell a sokasági értékösszeget, és
2. ezt össze kell vetni a kimutatásban szereplő összeggel.

A következőkben tehát tekintsük az Y_i és X_i ($i=1\dots N$) páros sokaságot, ahol Y_i jelöli az ismert, kimutatásban szereplő tételeket, és X_i ezek valós, de nem ismert értékét. Ez alapján a könyvvizsgáló feladata annak eldöntése egy előre adott (pl. 95%) bizonyossággal, hogy a teljes könyv szerinti érték ($Y = \sum_i Y_i$) és a valós érték ($X = \sum_i X_i$) különbsége hogyan viszonyul a lényegességi küszöbhez. Amennyiben az eltérés nem haladja meg a lényegességi küszöböt, elfogadja a kimutatást, ellenkező esetben elutasítja.⁶⁴

A jóhiszemű feltevés szerint a könyvvizsgáló minden általa megvizsgált Y_i esetén képes X_i pontos megadására, de mivel megelégszik a részleges bizonyossággal, ezért döntését az összesen N tételből n megvizsgálásával fogja meghozni. A mintába kerülő n tételt – az

⁶⁴ A könyvvizsgálat nem csak elfogadó és elutasító véleménnyel végződhet: létezik az ún. korlátozó záradék és a megbízástól történő ellátás intézménye is. Korlátozó záradékot ad például a könyvvizsgáló akkor, ha bizonyos megnevezett részterületeket leszámítva összességében megbízható a beszámoló. A könyvvizsgálónak jogában áll elállni a megbízástól, ha például olyan objektív akadály merül fel, mint a menedzsment negatív hozzáállása.

általános sokasági tételektől megkülönböztetendő – Y_i és X_i kisbetűs változataival (y_i és x_i) jelöljük. Adott tétel könyv szerinti és a valós értéke segítségével számíthatjuk a következő két mutatót:

- $d_i = y_i - x_i$ ($D_i = Y_i - X_i$), a minta (sokaság) i. elemében lévő hiba/eltérés (error vagy deviation), ami a minta esetében ismert, a sokaság általános elemére pedig nem ismert, de létező érték;
- $t_i = \frac{y_i - x_i}{y_i} = \frac{d_i}{y_i}$ ($T_i = \frac{Y_i - X_i}{Y_i} = \frac{D_i}{Y_i}$), a minta (sokaság) i. elemének szennyezettsége, tehát a könyv szerinti értékhez viszonyított relatív hibája (tainting).

Ha éles szemmel figyeljük, azonnal kitűnik a leírt modell legnagyobb hibája: ez a módszer nem alkalmas a „kifelejtett” tételek felderítésére. A továbbiakban tehát feltesszük, hogy nincsenek ilyen, „kifelejtett” tételek⁶⁵, és csupán a tételek értékelése lehet hibás.

A könyvvizsgálatban előforduló populációk főbb statisztikai jellemzői

Ahhoz, hogy megértsük, miért is problematikus a könyvvizsgálatban az értékösszeg becslése, mindenképpen be kell mutatni a populáció jellegzetességeit. Ezzel kapcsolatosan az 1960-as, az 1970-es és az 1980-as években sok tanulmány született, melyek fő eredményeit ez a szakasz foglalja össze.

Az első szembetűnő jelenség, hogy a tételek túlnyomó része helyes, azaz nem tartalmaz hibát. Ez azzal jár, hogy a megvizsgált mintának csak minimális része tartalmaz érdemi információt a hibákról. Johnson–Leitch–Neter [1981] tanulmányából kiderül, hogy az általuk vizsgált adatállományokban a „Vevők” tételek (a B/II. „Követelések” egy alcsoportja) hibaarányának mediánja 0,024 (a kvartilisek $Q_1=0,004$ és $Q_3=0,089$), míg a „Készletek” ellenőrzésekor ugyanezek a mutatók $Q_1=0,073$, $Q_2=0,154$ és $Q_3=0,399$. Ezért például a nagy számok törvénye szerint a vevőállományból vett 500 elemű (tehát nagy) mintánál körülbelül 12 darab tétel információtartalma alapján kell az egész sokaságról nyilatkozni. A tanulmány szerint a nem nulla hibák (eltérések) eloszlása lényegesen eltér az egyes beszámoló-területeken. Míg a vevők esetén például szinte kizárólag csak túlértékelések (overstatement) szerepelnek, addig a

⁶⁵ Ezt azért tehetjük fel, mert a könyvvizsgálat során a könyvvizsgáló egyéb módon már meggyőződött arról, hogy a szervezet belső eljárásai garantálják-e a kimutatások teljes körűségét. Amennyiben nem, akkor pedig úgyszemint fogja becsülni ilyen módon az értékösszeget.

készleteknél az alul- és túlértékelések körülbelül fele-fele arányban fordultak elő. Más tanulmány (Ham–Lassel–Smieliauskas [1985]) kitért a „Szállítók” tételeire is, ahol az alulértékelés (understatement) volt a tipikus.

A második specialitása ezeknek a sokaságoknak, hogy a nagyobb könyv szerinti értékű tételek nagyobb valószínűséggel tartalmaznak hibát, ám a relatív hiba (eltérés/könyv szerinti érték) nagysága nincs szignifikáns kapcsolatban a könyv szerinti értékkel. Ezen felül, az eltérés mértékének szórása a könyv szerinti értékkel növekszik. Amennyiben az adott tétel egyes pénzegységeit, pontosabban a tétel ezekhez rendelt relatív hibáját tekintjük sokaságnak, akkor a leírtak alapján látható, hogy ennek a sokaságnak jelentős része a 0 körül koncentrálódik. Emellett, például a vevőknél, megfigyelhető egy csomópont az 1 körül is, ugyanis a hibák jelentős része 100 százalék túlértékelés (például a már befolyt bevételt nem rendezték számvitelileg). A harmadik fontos probléma, hogy a legtöbb sokaság ferde, továbbá a ferdeség jellemző iránya és mértéke más és más az egyes beszámolóterületeken. A most felsorolt tulajdonságok miatt az általánosan használt eloszlások (normális, exponenciális, gamma, béta stb.) nem alkalmasak a valós és a könyv szerinti érték eltéréseinek modellezésére.

A valós érték pontbecslése

Tegyük fel, hogy n elemű egyszerű véletlen (a továbbiakban: EV-) mintát vettünk a sokaságból. Ez alapján egyebek mellett a következő módokon becsülhetjük a sokasági értékösszeget:

Legegyszerűbb a mintaátlag alapján történő becslés: $\hat{X}_m = N \cdot \frac{\sum x_i}{n}$. Ennek a becslésnek nyilvánvaló hátránya, hogy nem használja fel a pénzügyi kimutatásban szereplő, a valós értékkel jól korreláló adatokat.

A meglévő információt az $\hat{X}_d = Y - N\bar{d}$, az $\hat{X}_r = Y \cdot \left(\frac{\bar{x}}{\bar{y}} \right)$ különbség-, illetve hányadosbecsléssel, illetve a fenti három valamilyen súlyozott átlagával használhatjuk fel.

Végül tételezzük fel, hogy az EV minta helyett olyan visszatevés nélküli mintát vettünk, ahol minden sokasági elem mintába kerülési valószínűsége egyenesen arányos annak könyv szerinti értékével (a továbbiakban: PPS minta, az angol probability proportional to size

rövidítésből). Ilyen mintavételi terv mellett a sokasági értékösszegre adott torzítatlan Horvitz-

$$\text{Thompson becslés } \hat{X}_{HT} = \sum \frac{x_i}{\left(n \frac{y_i}{Y}\right)}.$$

A valós érték intervallumbecslése

Az ellenőrzési gyakorlatban alapvetően az egyoldali intervallumok terjedtek el, ezért a továbbiakban csak a $100 \cdot (1 - \alpha)$ százalékos felső határral fogunk foglalkozni.

A legegyszerűbb módja az intervallumbecslésnek a pontbecslés mintavételi eloszlását veszi alapul, nevezetesen annak első két (centrális) momentumát, tehát az átlagot és a varianciát. A tipikus becslési szituációkban tehát egy kétoldali intervallumbecslés a $\mu \pm \kappa_{1-\alpha/2} \cdot \sigma$ képlettel adható meg, ahol $\kappa_{1-\alpha/2}$ a pontbecslés standardizált eloszlásának megfelelő kvantilis értéke. Azokban az esetekben, amikor egy eloszlás „jól viselkedik”, a központi határeloszlás tétel alapján ezek a becslések már kisebb minták esetén is elfogadható eredményekre vezetnek.

A korábban leírt főbb statisztikai jellemzőkből kitűnik, hogy a könyvvizsgálat tipikus sokaságai nem követnek „jól viselkedő” eloszlást, és Neter et al. [1975, 1977] vizsgálatai kimutatták, hogy ezeknél a sokaságoknál a hagyományos intervallumbecslési módszerek valóban jelentősen torzítanak. A torzítás egy része a ferdeségből, másik része a kurtózisból adódik, melyeknek az a folyománya, hogy a $\frac{\text{pontbecslés} - \text{valós érték}}{\text{pontbecslés mintavételi szórása}}$ hányados még megközelítőleg sem követ t-eloszlást. (Kaplan, [1973a, 1973b])

Mivel a hagyományos módon készített becslések nem adtak kielégítő eredményt az eltérés nagyságára, a statisztika és a könyvvizsgálat határterületén több alternatív következtetési eljárást is kifejlesztettek, ezek egy része nagyban támaszkodik a sokasági aránybecslés módszereire.

A sokaság elemeiben található hiba eloszlását kevert eloszlással⁶⁶ modellezzük: a hiba p valószínűséggel egy ξ valószínűségi változó értékeit ($E(\xi) = \theta, \xi \neq 0$) veszi fel, $1-p$

⁶⁶ Kevertnek nevezzük egy olyan valószínűségi változó eloszlását, amelynek értékeit úgy származtatjuk k darab különböző, előre rögzített eloszlásból, hogy a k . valószínűségi változó értékét pontosan p_k valószínűséggel veszi fel.

valószínűséggel pedig 0^{67} . $p_{1-\alpha}(m, n)$ jelöli M/N sokasági arány $1-\alpha$ megbízhatóságú felső korlátját N elemű sokaság, n elemű minta, M minősített sokasági elem és m minősített mintabeli elem esetén. (lásd a III. részt)

Egy EV mintán alapuló becslés

Ez a becslés feltételezi, hogy:

- a hiba túlértékelésből fakad ($Y_i \geq X_i$, azaz $D_i \geq 0$);
- a kimutatott tételek mind pozitívak ($Y_i > 0$);
- a tételben lévő hiba maximális értéke legfeljebb a tétel értéke ($D_i \leq Y_i$, azaz $X_i \geq 0$).

Mindezeket figyelembe véve felírható a következő két reláció:

$$E(\xi) = \theta \leq Y_{\max}, \quad (4.1)$$

mivel ξ minden D_i realizációjára $D_i \leq Y_i \leq Y_{\max}$, illetve

$$E(D_i) = \sum_i \frac{D_i}{N} = p\theta \leq pY_{\max}, \quad (4.2)$$

hiszen D keverésében a nem-nulla változó valószínűsége p volt. Ebből egyszerű átalakítással kapjuk a sokasági hiba összértékére a $\sum D_i \leq NpY_{\max}$ felső korlátot. Amennyiben n elemű mintát veszünk a sokaságból, amelyben m hibás tételt találtunk, a

$$\hat{D}_{1-\alpha, EV} = Np_{1-\alpha}(m, n)Y_{\max} \quad (4.3)$$

becslésre fennáll a $\Pr\left(\sum D_i \leq \hat{D}_{1-\alpha, EV}\right) \geq \Pr\left(NpY_{\max} \leq \hat{D}_{1-\alpha, EV}\right) = 1-\alpha$ reláció, tehát a becslésünk *legalább* $100 \cdot (1-\alpha)\%$ -ban megbízható.

Figyeljük meg, hogy ez a becslés nem használja fel a mintában megfigyelt eltérések *nagyságát*, csupán a minta hibás tételeinek *arányát*, ezért *elvileg* jelentős pontosságjavulást lehet elérni egyrészt az Y értékei szerinti rétegzett mintavétellel, másrészt a maximális hibanagyságra tett feltevés módosításával. A gyakorlatban ennek ellenére ezt a módszert

⁶⁷ A kevert eloszlás definícióját figyelembe véve a sokaságban található hiba eloszlását 2 valószínűségi változóból „kevertük ki”: egy tetszőleges olyan ξ valószínűségi változóból, melyre $\xi(\omega) \neq 0 \quad \forall \omega$ esetén, és egy „determinisztikus” valószínűségi változóból, amely konstans 0 értékű.

ritkán használják, alapvetően azért, mert szinte kivétel nélkül PPS elvű (ezen belül is MUS) mintavételt alkalmaznak⁶⁸.

MUS mintán alapuló becslések

A beszámolók auditálásakor az egész világon széles körben használt, gyakorlatban előforduló becslési eljárások legfontosabb közös jellemzője a MUS-, vagy DUS-mintavétel (monetary unit sampling vagy dollar unit sampling, a hazánkban elterjedt terminológia szerint pénzegység alapú mintavétel). A MUS a könyvvizsgálói gyakorlatban olyannyira elterjedt és elfogadott módszer lett, hogy szinte mást nem is használnak, és általában figyelmen kívül hagyják a módszer meglévő korlátjait, előfeltevéseit, így sokszor azokra a következtetésekre is ezt alkalmazzák, amikre alkalmatlan.

A MUS valójában csak annyit jelent, hogy a mintát az eredeti sokaság pénzegységeiből alkotott mesterséges sokaságból veszik, majd megvizsgálják azokat az eredeti tételeket, amelyekből pénzegységet választottak⁶⁹. Könnyen bizonyítható, hogy ez a kiválasztási módszer az eredeti sokaságra nézve egy PPS mintát eredményez.

MUS mintavétel esetén a hipotetikus sokaság „tételszáma” (elemszáma) Y (az *eredeti sokasági értékösszeg*), a tételek (sokasági elemek) „könyv szerinti értéke” pedig definíció szerint a hipotetikus sokaság minden egyes elemére 1. A minta elemszámához (n) képest Y gyakorlatilag végtelennek tekinthető (párszáz vs. többmillió/milliárd), ezért mindegy, hogy visszatevéssel vagy visszatevés nélkül veszünk-e mintát.

Ennek a mintaválasztási megközelítésnek a könyvvizsgálatban való alkalmazására tett első utalás még 1961-ből, van Heerden holland nyelvű cikkéből származik, de a könyvvizsgálói szakma szélesebb köre csak 1963-ban ismerte meg, Kenneth W. Stringertől. Az általa akkor még csak nagy vonalakban leírt MUS módszer egyik legismertebb becslési eljárását Stringer-féle felső határnak (Stringer bound) hívják.

A hetvenes években több alternatív módszert is kifejlesztettek, melyek legtöbbje azonban továbbra is magán hordozza a Stringer-féle becslés gyengeségeit: a becslések torzítatlansága analitikusan nem igazolható, a szimulációk alapján pedig a becslések jó része túlságosan

⁶⁸ Vegyük azonban észre, hogy a PPS mintavétel a rétegzett mintavétel speciális határeseteként értelmezhető.

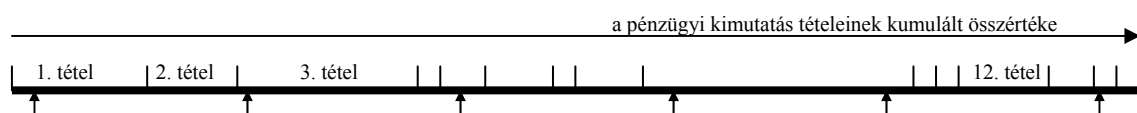
⁶⁹ A MUS a statisztikai mintavétel más területein ismert „kumulált értékösszegek módszere” egy alkalmazásának is tekinthető.

konzervatív, azaz a névleges szintnél jóval magasabb a megbízhatóságuk, és így jóval kisebb a pontosságuk (túlságosan széles az intervallum).

1. A MUS minta kiválasztásának technikai lebonyolítása

A pénzegység alapú mintavételt technikailag többféleképpen lehet elvégezni, melyből a következő módszereket érdemes kiemelni:

1. Az elmélet szempontjából legegyszerűbb esetben a pénzegységekből ún. korlátozás nélküli mintát, azaz EV mintát veszünk. Ebben az esetben akár az is előfordulhat, hogy mind az n alkalommal ugyanazt a tételt kell megvizsgálnunk, és emiatt a korlátozás nélküli MUS mintavétel sokak számára nem elfogadható. Ennek ellenére ezt a mintavételi technikát tekintjük alapértelmezésnek a továbbiakban számtalan jó tulajdonsága miatt.
2. A tételek valamilyen előre rögzített sorrendben kumulált sorozatának egy véletlenszerűen kiválasztott pontjáról elindulva n alkalommal felmérünk Y/n nagyságú lépésközt. Ez a manuális gyakorlatban leginkább elterjedt módszer, a szisztematikus kiválasztási módszerekre jellemző egyszerűségének köszönhetően. A könnyebb megértés kedvéért tekintsük a következő ábrát:

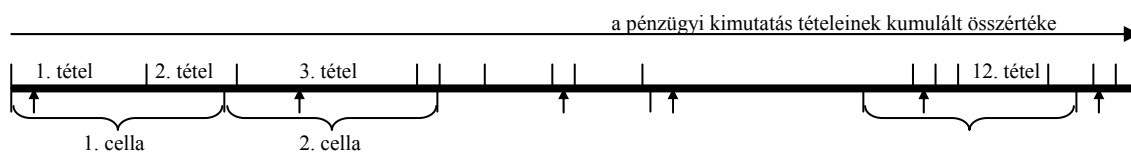


9. ábra: Tételek a szisztematikusan kiválasztott MUS mintában

A rajzon a felső beosztás intervallumai jelzik az egyes tételeket. Az első nyíl mutatja a véletlenszerűen kiválasztott pontot, a két szomszédos nyíl közötti távolság pedig a lépésközt. Vegyük észre, hogy a fenti rajzon a 4. és az 5. nyíl ugyanazt a tételt jelöli meg: Az ilyen tételeket nevezik lépésköz feletti, vagy nagy értékű tételeknek (HVI – high value items). A MUS egyes változatai a nagy értékű tételeket más és más módon kezelik, de a mi szempontunkból ez most nem lényegi kérdés.

3. Az úgynevezett cella-módszerben a 2. pontban leírttal szemben a tételek kumulált sorozatát n db egyenlő, Y/n hosszúságú intervallumra („cellára”) osztjuk, és minden egyes intervallumon belül véletlenszerűen kiválasztott 1-1 pont (az ábrán továbbra is

nyíllal jelölve) határozza meg a mintaelemeket. A módszerhez külön kiértékelő formula is tartozik, aminek részleteiről külön alponthan fogunk írni.



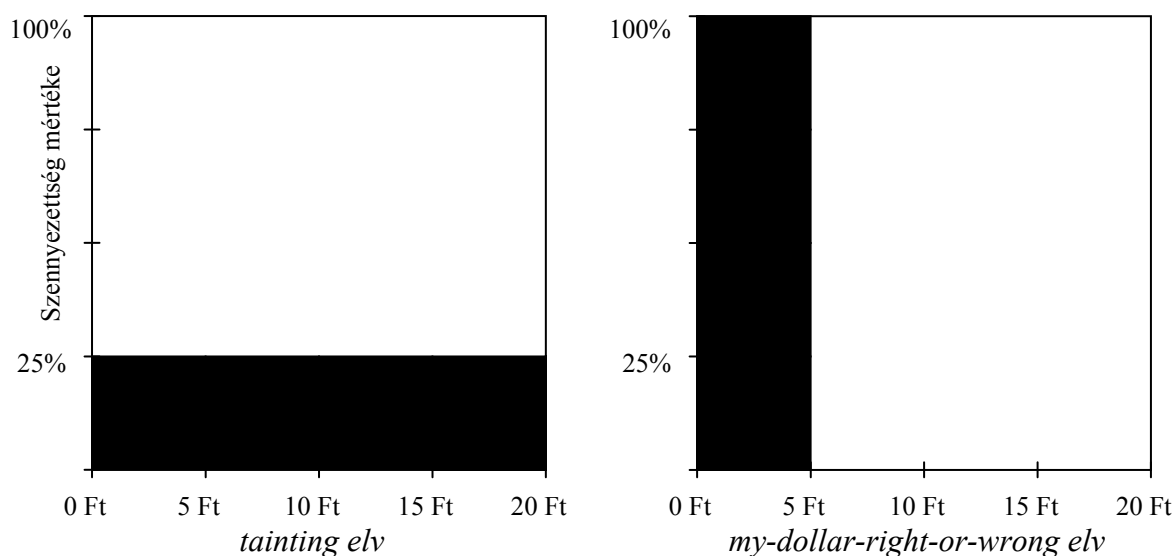
10. ábra: Cellás módszerrel kiválasztott tételek a MUS mintában

A fenti három mintavételi terv a tételek rögzített sorrendje esetén nem egyenértékű. Vegyük észre, hogy sem a 2., sem a 3. terv nem képes produkálni minden olyan mintát, amit az 1. módszer eredményezhet: sem a 2., sem a 3. mintavételi terv nem tud például olyan mintát eredményezni, amiben mind a 6. mind a 7. tétel szerepel. Hasonló módon a 2. terv sem képes minden olyan mintát produkálni, amit a 3. tud. Könnyen látható azonban, hogy a tételek mintavétel előtti „megkeverésével” (véletlenszerű permutálásával) ez a különbség megszűnik. Nehéz analitikusan átlátni, hogy a mintavételi terveknek ez a különbsége pontosan milyen hatást gyakorol egy adott kiértékelő modulra, és ezzel kapcsolatosan az általam ismert irodalom sem nyújtott kellő mértékű eligazítást.

További problémát okoz ezeknél a mintavételi terveknél annak eldöntése, hogy a mintába választott különböző pénzegységek milyen mértékű hibát tartalmaznak. Ezzel kapcsolatosan két felfogás létezik. Az uralkodó, de időben későbbi megközelítés (tainting-elv) szerint a pénzegység hibája az őt tartalmazó fizikai tétel szennyezettségével egyezik meg, tehát bárhonnét vesszük ki az adott tételből a mintaelemet, a hiba ugyanaz. A másik megközelítés (my-dollar-right-or-wrong) az adott fizikai tétel hibáját a tétel elejétől (egyes alkalmazásokban a végétől) kezdi számolni, tehát attól függően, hogy honnét származik a mintaelem, a hibája 1 vagy 0^{70} . Ha tehát a példa kedvéért egy tétel 25 százalékos szennyezettségű, és 20 pénzegységből áll, akkor a tainting-elv szerint minden egyes pénzegység 25 százalékos szennyezettségű. Ezzel szemben a másik megközelítésben az első 25 százalékot (az első 5 egységet) 100 százalékosan szennyezettnek, a továbbiakat viszont teljesen szennyezettségmentesnek tekintik⁷¹. Ezt a helyzetet a 11. ábra szemlélteti.

⁷⁰ Vagy pedig egy tört, ha pont a határon van, és az eredeti tételben szereplő hiba nem egész szám forintban nézve. De ez az eset általában irreleváns, mivel a forintra kerekítés nem sok hibát okoz várható értékben.

⁷¹ A my-dollar-right-or-wrong megközelítés gyakorlatilag nem más, mint a sokasági arány becslés közvetlen alkalmazása a pénzegységek mesterséges populációjára.



11. ábra: a szennyezettség mérése (könyv szerinti érték 20 Ft, valós érték 15 Ft)

Érdekes módon a tainting megközelítés – annak ellenére, hogy kismintás tulajdonságai jobbak, aszimptotikusan alulmarad a my-dollar-right-or-wrong megközelítéssel szemben (lásd például *Pap–van Zuiljen* [2000]). A továbbiakban mi a tainting megközelítést fogjuk alkalmazni. A mintaválasztási módok és a hibák különböző értékelései közötti eltérések jobb megértésére egy leegyszerűsített példán kövessük végig alkalmazásukat.

Először tekintsük a következő elszámolást, amit egy külföldi kiküldetésből hazatért kolléga nyújtott be.

Sorszám	Megnevezés	Könyv szerinti érték (Y_i)	Valós érték (X_i) /a priori ismeretlen/	Megjegyzés	Relatív hiba
1	Taxiszámla Ferihegyre	23 €	0 €	Kapott BKV bérletet, így nem jogosult elszámolni	100 %
2	Repülőjegy oda vissza	512 €	512 €		0 %
3	Taxi a szállásig	72 €	72 €		0 %
4	Szállás 5 napra félpanz.	432 €	324 €	Elírás	25 %
5	Helyi tömegk. heti bérlet	84 €	84 €		0 %
6	Éttermi ebéd 1. nap	15 €	15 €		0 %
7	Éttermi ebéd 2. nap	15 €	15 €		0 %
8	Éttermi ebéd 3. nap	15 €	15 €		0 %
9	Éttermi ebéd 4. nap	15 €	15 €		0 %
10	Éttermi ebéd 5. nap	15 €	15 €		0 %
11	Telefonköltség	43 €	43 €		0 %
12	Taxi a repülőtérre	68 €	68 €		0 %
13	Taxi Ferihegyről	35 €	0 €	lásd. 1. tételnél	100%
Összesen		1344 €	1178 €		12 %

12. Táblázat: egy hipotetikus útiköltség-elszámolás

A tételek közül szűrőpróbaszerűen kiválaszt a pénzügyes 6 tételt. Ehhez először el kell készíteni a tételek kumulált sorozatát:

Sorszám	Könyv szerinti érték	Kumulált könyv szerinti érték
1	23 €	23 €
2	512 €	535 €
3	72 €	607 €
4	432 €	1039 €
5	84 €	1123 €
6	15 €	1138 €
7	15 €	1153 €
8	15 €	1168 €
9	15 €	1183 €
10	15 €	1198 €
11	43 €	1241 €
12	68 €	1309 €
13	35 €	1344 €

13. Táblázat: az útiköltség-elszámolás tételeinek könyv szerinti- és kumulált könyv szerinti értéke

- Az első mintavételi módszer alkalmazásához 6 elemű EV mintát veszünk az 1, 2, ... 1344 számokból. Legyenek ezek a 17, 52, 364, 836, 1293 és 1317. Ezt felhasználva a mintánk az 1., 2., 2., 4., 12. és 13. tételekből áll. Látható, hogy a 2. tétel kétszer szerepel a mintában.
- A második mintavételi módszerhez két értéket kell meghatározni: a lépésközt és a kezdőpontot. A lépésköz Y/n , azaz $1344/6=224$, a kezdőpont pedig az 1, 2, ... 1344 számokból választott 1 elemű EV minta, ami ez esetben legyen mondjuk 3. A mintába eső pénzegységek a 3, $3+224=227$, $3+224+224=451$, 675, 899, 1123. Fizikai tételekre lefordítva ez az 1., 2., 2., 4., 4., 5. tétel. A 2. és a 4. tétel HVI, ezért mindenképpen a mintába kellett kerülniük. A 2. tétel nagysága a lépésköz kétszeresét is meghaladja, ezért mindenképpen kétszer kerül a mintába. A 4. tétel bizonyos kezdőpontok esetén egyszer, bizonyos kezdőpontok esetén kétszer kerül a mintába.
- A cella módszerhez először meg kell határozni a cellákat. A cellák nagysága megegyezik a lépésközzel, így a cellahatárok 1-224, 225-448, 449-672, 673-896, 897-1120, 1121-1344. Ezek után 6 elemű visszatevéses mintát veszünk az 1-224 intervallumból, legyen ez 27, 143, 53, 197, 81, 152. A kiválasztott pénzegységek az $1+27-1=27$, $225+143-1=367$, $449+53-1=501$, $673+197-1=869$, $897+81-1=977$, $1121+152-1=1272$. Az ezekhez tartozó tételek sorszáma rendre 2, 2, 2, 4, 4, 12. A cella módszerben a szerencsétlen véletlenek miatt egy tétel annyiszor belekerülhet a mintába, ahány cellával van metszéspontja. Ezeket az ismétléseket a cella módszer egy fejlettebb megvalósítása kiszűri, de ennek ismertetésétől most eltekintek.

A két hibamérési megközelítés közötti különbséget a második módszerrel választott mintán mutatjuk meg, konkrétan a 4. elszámolt tétel esetén. Emlékezzünk, hogy ez a tétel kétszer került a mintába. A 4. tétel a 608. pénzegységtől az 1039. pénzegységig tartott.

- Az első felfogás, a tainting elv szerint a mintába került 432 euróból 75%-nyi helyes, 25%-nyi hibás, mindkét esetben; $(324/432=0,75=75\%)$
- A hibát a tétel elejétől felmérő my-dollar-right-or-wrong felfogás szerint a hibás pénzegységek a 608-tól 715-ig tartanak $((1039-608)*0,25+608=715)$. Így a mintába került első pénzegység (675) hibája 100% (mert $675<715$), a második pénzegység (899) hibája 0% (hiszen $899>715$);
- A hibát a tétel végétől felmérő my-dollar-right-or-wrong felfogás szerint a 931-től 1039-ig található pénzegységek a hibásak. Így a mintába került mindkét pénzegység 0% hibát tartalmaz.

2. Az EV becslés alkalmazása a mesterséges sokaságra

Az EV mintán alapuló (4.3) becslést alkalmazva a MUS mesterséges sokaságára a következőt kapjuk (Y db tétel, a tételek „könyv szerinti értéke” pedig definíció szerint 1):

$$\Pr\left(\sum D_i \leq Y p_{1-\alpha}(m, n)\right) \geq 1 - \alpha. \quad (4.4)$$

Mivel az eredeti $Y \leq N Y_{\max}$, ezért a MUS mintán alapuló

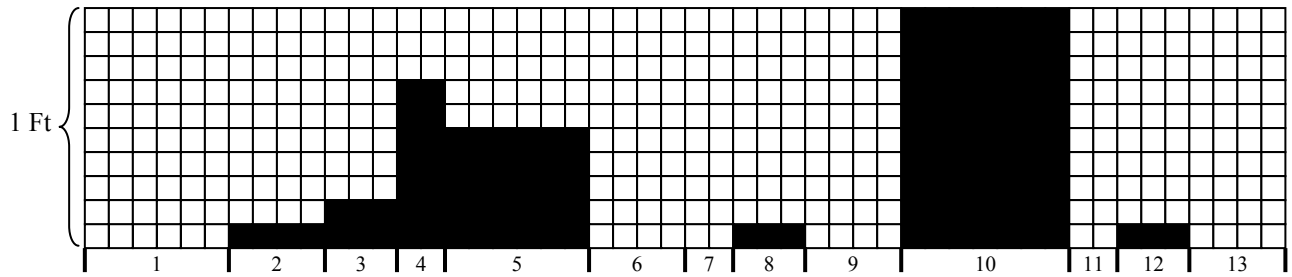
$$\hat{D}_{1-\alpha, \text{MUS}} = Y p_{1-\alpha}(m, n) \quad (4.5)$$

becslés jóval pontosabb (kevésbé konzervatív, szűkebb az intervallum) az EV mintán alapuló becslésnél. Mindazonáltal ez a becslés sem veszi figyelembe, hogy nem minden hibás tétel 100%-ig hibás, így ez a becslés is túlságosan óvatosnak tekinthető.

Az eddig leírt két elemi módszernek minden hibájuk ellenére megvolt az a jó tulajdonsága, hogy a megbízhatóságuk analitikusan igazolható. A most bemutatandó becslésekre ez már sajnos nem, vagy csak korlátozottan lesz igaz. Ezeknél a kapcsolódó irodalom szinte kivétel nélkül szimulációkkal igyekszik a megbízhatóságról, illetve a torzítás mértékéről meggyőződni.

3. A Stringer-féle felső határ (Stringer bound)

Mielőtt leírnánk a Stringer becslést részleteiben, tekintsük meg az alábbi ábrát:



12. ábra: A könyv szerinti érték szennyezettsége

Az ábrán látható kis négyzet magassága 0,1 Ft, szélessége 1 Ft, tehát a rács egy oszlopa jelenti a sokaság egy elemét (azaz 1 forintnyi pénzegységet). Az eredeti sokaságot az alsó, vastagabb vonalkák definiálják, tehát az eredeti sokaság könyv szerinti értékei rendre 6, 4, 3, 2, 6... stb. Mivel ez esetben feltételezzük, hogy a valós érték nem negatív, de legfeljebb a könyv szerinti érték, ezért egy adott tétel valós értékét úgy ábrázoljuk, hogy befeketítjük a könyv szerinti érték szennyezettség mértékének megfelelő hányadát. Ez alapján pl. a második tétel könyv szerinti értéke 4, valós értéke 3,6; a 10. tétel pedig nem ér semmit a valóságban.

A korábban leírt (4.5) becslést felírhatjuk a következő alakban is:

$$\hat{D}_{1-\alpha, \text{MUS}} = Y_{p_{1-\alpha}}(m, n) = 10 \cdot 0,1 \cdot Y_{p_{1-\alpha}}(m, n) = \sum_{i=1}^{10} 0,1 \cdot Y_{p_{1-\alpha}}(m, n) \quad (4.6)$$

A 12. ábra által felvázolt helyzetben ez azt jelenti, hogy külön számoltuk minden 10 fillérre felső korlátot, amiket aztán összeadtunk. Ha pontosítani szeretnénk a becslést, az első intuíció ebből a képletből kiindulva azt sugallhatja, hogy nézzük meg, hány mintába került tételben van *legalább* 1, 2, ... 100 fillér hiba (ez nyilván egy monoton csökkenő sorozat lesz, ugyanis ha $a > b$, akkor a legalább b hibát tartalmazó tételek halmaza tartalmazni fogja a legalább a hibát tartalmazó tételek halmazát. A két halmaz különbsége a pontosan $b+1$, $b+2$, $b+3$... a mennyiségű hibát tartalmazó tételek uniója). Ezután külön-külön adjunk felső becslést az adott kategória (tehát például legalább 2 fillér hiba stb.) sokasági arányára, és ezeket a kategóriák nagyságával (ebben az esetben $Y/100$) súlyozva adjuk össze. Amint nemsokára látni fogjuk, ez egy, a Stringer-féle felső határhoz nagyon hasonló felső határt fog megadni. A gondolatmenetben azonban egy súlyos hiba van: nem igaz ugyanis, hogy $\Pr(\xi \leq x) = 1 - \alpha$ és $\Pr(\eta \leq y) = 1 - \alpha$ relációkból következne $\Pr(\xi + \eta \leq x + y) = 1 - \alpha$ reláció, csupán

$\Pr(\xi + \eta < x + y) \in [1 - 2\alpha; 1]$ állítható. A Stringer-féle felső határ egy rendezett mintás statisztika: a nagyság szerint csökkenő szennyezettségeket rögzített (csak a mintamérettől függő), a helyezésüknek megfelelő súlyokkal „átlagoljuk”, majd ezt az „átlagot” megszorozzuk a teljes sokaság könyv szerinti értékével. Képlettel:

$$\hat{D}_{1-\alpha, st} = Y \cdot \left(p_{1-\alpha}(0, n) \cdot 1 + \sum_{i=1}^n [p_{1-\alpha}(i, n) - p_{1-\alpha}(i-1, n)] \cdot t_i \right) \quad (4.7)$$

ahol t_i a mintában található i . legnagyobb szennyezettséget jelenti. Ezt a képletet átalakítva kapjuk:

$$\begin{aligned} \hat{D}_{1-\alpha, st} &= Y \cdot \left(p_{1-\alpha}(0, n) \cdot 1 + \sum_{i=1}^n [p_{1-\alpha}(i, n) - p_{1-\alpha}(i-1, n)] \cdot t_i \right) = \\ &= Y \cdot \sum_{i=0}^n (t_i - t_{i+1}) \cdot p_{1-\alpha}(i, n) \end{aligned} \quad (4.8)$$

ahol $t_0 = 1$ illetve $t_{n+1} = 0$.

A 12. ábra példájában 100, 70, 50, 20, 10 és 0 százalékos szennyezettségeket látunk, tehát ez a képlet rendre felső becslést ad a legalább 100, 70, 50, 20, 10 százalékos hibák arányára, ezeket 100-70, 70-50, 50-20, 20-10, 10-0 százalékos súlyokkal súlyozza, majd ezt a súlyozott összeget (átlagos hibamértéket) kiveti a teljes sokaságra, Y -ra. Ha ebben a formában nézzük a Stringer-féle felső határt, azonnal látszik, hogy ebben az esetben is az intervallumokba esés valószínűségének felső határát becsültük meg, és elkövettük azt a fent említett hibát, hogy ezeket az értékeket mechanikusan összeadtuk. Természetesen ez még nem jelenti azt, hogy a Stringer becslés nem lenne jó, de ezzel a megközelítéssel nem bizonyítható a jósága.

Amikor az eredeti képlet megjelent, még nem állt olyan számítástechnikai háttér rendelkezésre, mellyel gyorsan meghatározható lett volna $p_{1-\alpha}(m, n)$ értéke tetszőleges N , m , n és megbízhatósági szint esetén. Mivel nagyobb sokaságokra a hipergeometriai eloszlás közelíthető binomiálissal, illetve alacsony hibaarányok mellett a binomiális eloszlás közelíthető a könnyen kezelhető Poisson eloszlással, ezért kezdetben a $p_{1-\alpha}(m, n)$ érték helyett a $\frac{\lambda_{1-\alpha}(m)}{n}$ közelítő értéket használták, ahol $\lambda_{1-\alpha}(m)$ a Poisson eloszlás paraméterére vonatkozó $100 \cdot (1 - \alpha)\%$ -os felső határ m megfigyelt hiba mellett. Ennek a közelítésnek megvolt az az előnye, hogy a rendezett mintás statisztika súlyait táblázatba lehetett gyűjteni a mintamérettől függetlenül, hiszen az kiemelhető volt a képletből:

$$\hat{D}_{st} = \frac{Y}{n} \cdot \left(P_0 + \sum_{i=1}^n P_i \cdot t_i \right), \text{ ahol } \frac{Y}{n} \text{ a lépésköz (itt vesszük figyelembe a mintaméretet), } P_i$$

pedig a táblázatból kiolvasható i . úgynevezett Poisson-faktor, tehát $P_i = \lambda(i) - \lambda(i-1)$. (Az irodalom nem teljesen következetes, egyes szerzőknél $\lambda(i)$ -t nevezik Poisson-faktornak.) Valószínűleg egyébként éppen ezért a könnyű kezelhetőségért terjedt el ebben a formájában a képlet, és nem a másik, $p_{1-\alpha}(m, n)$ szerint csoportosított formában. Noha ma már bármely korszerű személyi számítógép azonnal ki tudná számolni a hipergeometriai faktorokat is, a Poisson közelítéssel való számolás még mindig nagyon elterjedt a könyvvizsgálók között.

Mivel a Stringer-sejtés (tehát hogy a becslőfüggvény *legalább* $100 \cdot (1-\alpha)\%$ -ban megbízható) általános feltételek mellett mindezidáig nem került igazolásra, és ellenpéldát sem sikerült konstruálni, számos szimulációt végeztek és publikáltak a témában. A szimulációk erős empirikus bizonyítékot szolgáltatottak arra, hogy a becslés megbízhatósága jóval meghaladja a névleges $100 \cdot (1-\alpha)\%$ -ot (pl. az általánosan használt 95%-nál az esetek 98-99%-ban haladta meg a becsült felső határ a valós értéket).

Függetlenül azonban attól, hogy mekkora a becslés megbízhatósága, fennáll a $\hat{D}_{1-\alpha, st} \leq \hat{D}_{1-\alpha, MUS}$ reláció minden olyan esetben, amikor a szennyezettségek 0 és 1 közé esnek, ugyanis:

$$\hat{D}_{1-\alpha, st} = Y \cdot \sum_{i=0}^n (t_i - t_{i+1}) \cdot p_{1-\alpha}(i, n) \leq Y \cdot \sum_{i=0}^n (t_i - t_{i+1}) \cdot p_{1-\alpha}(m, n) = \hat{D}_{1-\alpha, MUS}, \quad (4.9)$$

mivel $t_0 = 1$ és m az a legkisebb egész szám, amelyre $t_m \leq 0$ (a szennyezettségek monoton csökkennek és nemnegatívak). Egyenlőség áll fenn, ha a szennyezettségek csak 0 vagy 1 értéket vehetnek fel. Ennek alapján kijelenthető, hogy amennyiben *legalább* $100 \cdot (1-\alpha)\%$ -ban megbízható a (4.7) becslés, akkor a pontossága jobb, mint a (4.5) becslésnek.

Ha a Stringer-féle felső határ (4.7) alatti alakjából (illetve az alakhoz kapcsolódó intuícióból) közelítünk, akkor a szumma első tagja veszi számba a mintában ugyan nem található, de feltételezhetően meglévő szennyezettséget. Az eredeti képletben ez $t_0 = 1$, azonban számos területen a gyakorlati tapasztalatok szerint jóval kisebb az elképzelhető legnagyobb hiba. Amennyiben tehát biztos információval rendelkezünk az elképzelhető legnagyobb hiba nagyságáról, akkor ezt az értéket rakva t_0 helyébe jelentősen élesíthetünk a becslésünkön (az így kapott becslés neve: generalized Stringer bound, tehát általánosított Stringer-féle felső

határ). A legfontosabb könyvvizsgálati szoftverekben is állítható ez az érték, általában BPP (basic precision pricing) a neve. Az elnevezés onnét származik, hogy nagyságrendileg általában ez a paraméter befolyásolja leginkább a becslésünket, és nem a mintából származó szennyezettségek.

Noha a Stringer-sejtést teljes egészében eddig nem igazolták, több fontos részeredmény született.

Az első, úttörőnek tekinthető írás Bickel [1992] tanulmánya. A szerző bizonyítja, hogy ha:

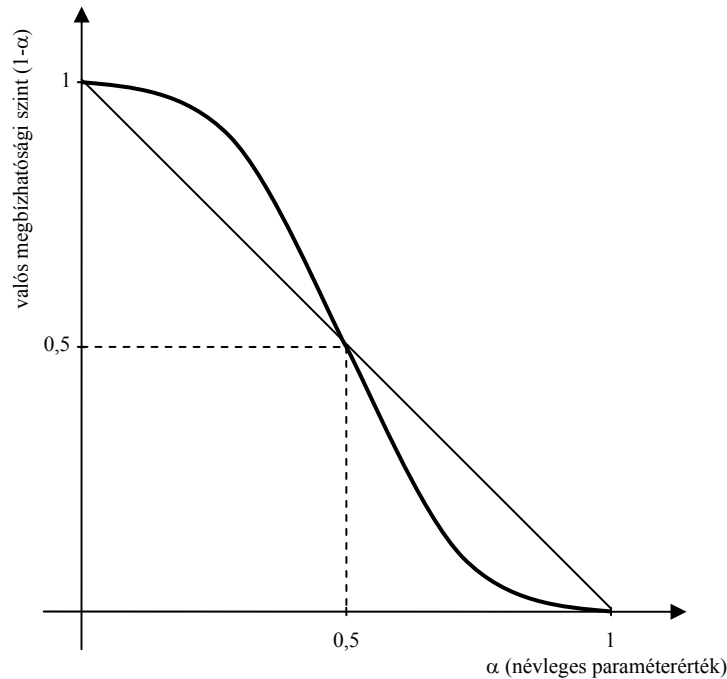
- ξ folytonos valószínűségi változó, akkor $P(D \leq \hat{D}_{1-\alpha, st}) \geq (1-\alpha)^{n+1}$, illetve ha
- ξ legfeljebb 2 értéket vehet fel, akkor $P(D \leq \hat{D}_{1-\alpha, st}) \geq 1-\alpha$, tehát a becslés legalább $100 \cdot (1-\alpha)\%$ -ban megbízható.

A cikk további fontos eredménye, hogy a becsült felső korlátot fel tudja írni ξ várható értékének, illetve az eloszlásfüggvény egy bonyolult integráljának összegeként. Ennek a felírásnak a jelentősége, hogy segítségével kiszámítható a becslés aszimptotikus (végtelen mintaméretnél értelmezett) eloszlása.

Pap, van Zuijlen és de Jager több tanulmányban [1995, 1996, 1997] folytatja a Bickel által elkezdett megközelítést. Három legfontosabb eredményük a következő:

- A Bickel által felírt aszimptotikus eloszlás segítségével bizonyítják, hogy $\alpha \in [0; 0,5]$ esetben a becslés aszimptotikusan megbízható, ellenkező esetben aszimptotikusan nem megbízható (lásd 13. ábra). A valós megbízhatósági szint az első esetben jóval meghaladja, a második esetben viszont még közelítőleg sem éri el a névleges $100 \cdot (1-\alpha)\%$ -ot. Mivel az auditorok általában 50% feletti megbízhatósági szinttel dolgoznak, ezért csak az $\alpha \in [0; 0,5]$ esetre tett megállapításoknak van gyakorlati jelentősége.
- Az előbbi észrevételt kiegészítve bevezetnek egy olyan módosított becselőfüggvényt, amely aszimptotikusan pontosan a névleges szintnek megfelelő megbízhatóságú.
- A szennyezettségek tetszőleges olyan eloszlása esetén, ahol a szennyezettségek csak 0 és 1 értéket vehetnek fel, minden olyan lehetséges $a_{1-\alpha}(i, n) \geq a_{1-\alpha}(i-1, n)$ együttható-

sorozatra, amelyre a $\hat{D}_{1-\alpha} = Y \cdot \left(a_{1-\alpha}(0, n) \cdot 1 + \sum_{i=1}^n [a_{1-\alpha}(i, n) - a_{1-\alpha}(i-1, n)] \cdot t_i \right)$ felső határ legalább $100 \cdot (1-\alpha)\%$ -ban megbízható, $a_{1-\alpha}(i, n) \geq p_{1-\alpha}(i, n) \quad \forall i$ -re. Ez azt jelenti, hogy ebben az esetben bizonyos tekintetben minimálisak a képletben szereplő együtthatók.



Forrás: Pap-van Zuijlen, [1996]

13. ábra: A Stringer-féle felső határ aszimptotikus megbízhatósága

A Stringer-féle felső határral kapcsolatos irodalom gyakorlati szempontból óriási jelentőségű cikke Neter et al. [1984] tanulmánya. A gyakorlatban ugyanis sokszor olyan nagy az auditálandó beszámoló, hogy azt csak részterületre bontva lehet vizsgálni. Az egyes részterületeken egymástól függetlenül történik a mintavétel, ennek ellenére véleményt kell mondani a teljes beszámoló megbízhatóságáról is. A legnagyobb problémát az jelenti, hogy a részterületek különbözősége miatt a teljes beszámolóra nézve már nem áll fent, hogy a mintába kerülés valószínűsége minden tételre arányos lenne a tétel beszámolóban szereplő nagyságával.

A cikk feltételezi, hogy az egyes részpopulációkból független MUS mintát vettek, amiket a Stringer-féle felső határt használva értékelték ki. A szerzők a kombinált felső határ kiszámolására 5 különböző megoldást javasolnak, melyek a következők:

1. Független valószínűségi változók összeadása.

Feltételezve, hogy az egyes részterületekre számított felső határok függetlenek egymástól, a teljes beszámoló felső határának megbízhatósági szintje legalább $\prod_j (1 - \alpha_j)$, ahol $1 - \alpha_j$ az j . részsokaságra tett felső határ becslés megbízhatósága. Ez alapján ha minden részsokaságra egyforma megbízhatósági szintet használunk, az aggregált becsléshez elvárt legalább $1 - \alpha$ megbízhatóság a részsokaságoknál legalább $(1 - \alpha)^{1/L}$ megbízhatóságú becsléseket igényel, ahol L a részsokaságok száma. (Tehát 95%-os megbízhatósághoz 2 részterület esetén mindkét részterületen 97,5%-os megbízhatóságú becslést kell készíteni.

2. Implicit standard hiba használata.

Mivel egy adott mintánál meghatározható nem csak az adott megbízhatóságú egyoldali intervallum végpontja, hanem a pontbecslés is, ezért implicit módon, „visszafelé”, kiszámíthatjuk azt a standard hibát, amit a normális eloszlással való közelítés esetén használva ugyanezt a felső határt kaptuk volna⁷². Formálisan:

$$SE = \frac{\hat{D}_{1-\alpha} - \hat{D}}{Z_{1-\alpha}}. \text{ Ezt a standard hibát használva kiszámolhatjuk a független változók}$$

összegének standard hibáját is (a négyzetösszeg gyöke), amiből a megszokott módon (pontbecslés + $z \cdot SE$) kapjuk az összegre vonatkozó becsült felső határt.

3. Közelítő globális kiértékelés.

$$\text{A felső határ becslés } \hat{D}_{1-\alpha, st} = Y_1 \cdot \left(p_{1-\alpha}(0, n_1) \cdot 1 + \sum_{i=1}^n [p_{1-\alpha}(i, n_i) - p_{1-\alpha}(i-1, n_i)] \cdot t_i \right)$$

képletében részsokaságonként különböző sokasági értékek (Y) és mintanagyságok (n) szerepelhetnek. Közelítő globális kiértékelés esetén a zárójelben szereplő első tag (az ún. basic precision) nélkül vesznek figyelembe minden sokaságot, kivéve azt a sokaságot, amelynél a legnagyobb az $Y_1 \cdot p_{1-\alpha}(0, n_1)$ szorzat értéke. Képlettel:

$$\hat{D}_{1-\alpha, st}^{kombinált} = \max_i (Y_i \cdot p_{1-\alpha}(0, n_i)) + \sum_i (\hat{D}_{1-\alpha, st, i} - Y_i \cdot p_{1-\alpha}(0, n_i))$$

⁷² A normális eloszlás esetén a becsült felső határ felírható „felső határ = pontbecslés + $z \cdot$ standard hiba” alakban. Ebből a standard hiba egyszerű átrendezéssel megkapható.

4. Globális kiértékelés.

A globális kiértékelés egy nagy mintának kezeli a részsokaságokból származó L különböző mintát, és erre a mintára alkalmazza a Stringer becslés egy módosított változatát (az „összefésült mintaelemeket itt is nagyság szerint csökkenő sorrendbe rakjuk):

$$\hat{D}_{1-\alpha, st}^{kombinált} = \max_i (Y_i \cdot p_{1-\alpha}(0, n_i)) + \sum_i [p_{1-\alpha}(i, n_i) - p_{1-\alpha}(i-1, n_i)] \cdot t_i \cdot Y_i, \text{ ahol az } i \text{ alsó}$$

index annak a mintának a nagyságára és kimutatott értékére utal, amiből az adott szennyezettség származik.

5. „Konzervatív” globális kiértékelés

A konzervatív változat annak a részmintának az n és Y paramétereit használja mindenhol, amelyre $Y_i \cdot p_{1-\alpha}(0, n_i)$ felveszi a maximumát. Azzal a kérdéssel nem foglalkozik a cikk, hogy ha több ilyen részminta is van, ugyanis akkoriban még az elméleti munkákban is a Poisson-közelítést használták, és a közelítő felírásban ennek a kérdésnek nincs jelentősége.

A tanulmányban szereplő 5 kiértékelési módszerre vonatkozóan több szimulációt is végeztek a szerzők, melyek kivétel nélkül 98-99%-os megbízhatóságot mutattak a kombinált felső határra 95% elvárt megbízhatóság mellett.

4. Cella-módszer (Cell bound)

A 2.1 szakasz 3. pontjában leírt mintaválasztási módszerhez a szerzők (Leslie, Teitlebaum, Anderson) külön kiértékelési módszert dolgozott ki. Mivel a módszer jóval kevésbé intuitív, mint akár a Stringer-féle felső határ, akár következő szakaszban ismertetésre kerülő multinomiális felső határ, ezért most csupán a kiértékelés módját írjuk le, intuitív indoklás nélkül. A most következő leírás megegyezik az eredetileg leírtakkal, így a Poisson eloszlással közelíti a valószínűségeket. Az elmúlt években a legtöbb gyakorlati alkalmazás (többek között az IDEA szoftver is) már az egzakt hipergeometriai faktorokat használja.

A mintán megfigyelt szennyezettségeket ebben az esetben is csökkenő sorrendbe rakjuk, és ezután alkalmazzuk a következő rekurzív formulát:

$$F(0) = \lambda_{1-\alpha}(0),$$

$F(i) = \max(F(i-1) + t_i, \lambda_{1-\alpha}(i) \cdot t_i)$, egészen az utolsó hibáig, m -ig.

A becsült felső határ a legutolsó F és lépésköz, Y/n szorzata:

$$\hat{D}_{1-\alpha, \text{cell}} = F(m) \cdot \frac{Y}{n}. \quad (4.10)$$

A cella-módszer kiértékelő része a szimulációk alapján kevésbé konzervatív, mint a Stringer-féle felső határ, azonban még így is jóval a névlegest meghaladó a megbízhatósága. Előnye, hogy a Poisson faktorokat tartalmazó táblázat segítségével számítógép nélkül is meghatározható.

5. Multinomiális felső határ (Multinomial bound)

A multinomiális módszert Fienberg, Neter és Leitch 1977-ben publikálta (Fienberg–Neter–Leitch [1977]), tehát gyakorlatilag egy időben került kifejlesztésre a Leslie, Teitlebaum, Anderson-féle cellamódszerrel (Leslie–Teitlebaum–Anderson [1980]). A cellamódszerrel szemben ez a becslés csak számítógép segítségével alkalmazható.

A multinomiális modell eredetileg leírt változatában minden pénzegységet besorolnak 101 kategória valamelyikébe aszerint, hogy az adott pénzegységre jutó szennyezettség mértéke 0%, 0%-nál több de legfeljebb 1%, 1%-nál több de legfeljebb 2% és így tovább 100%-ig. Ha

a sokaságban az i . csoportba eső elemek aránya p_i , akkor a sokasági hibaarányra $\sum_{i=0}^{100} \frac{i}{100} \cdot p_i$

felső becslést ad, és a becslés legfeljebb 1 százalékponttal haladja meg a valós értéket.

Amennyiben a mintába kerülő pénzegységeket visszatevéssel választjuk, vagy pedig a minta mérete a sokasághoz képest elhanyagolható, akkor a mintaelemek 101 kategória közötti eloszlása (n, p_i) paraméterű multinomiális eloszlást követ, ahol az első paraméter a mintaméretet, a további 101 db p_i paraméter pedig az indexe által meghatározott csoportba való esés valószínűsége. A multinomiális felső határ megadásához két lépés vezet.

Az első lépésben bevezetünk egy relációt a lehetséges minták között, melynek segítségével meghatározhatók a kapott mintánál „extrémebb” (azaz bizonyos kritériumok alapján kevesebb hibát tartalmazó) lehetséges minták. Nevezzük ezeknek a lehetséges mintáknak a halmazát S -nek! Mivel S -et alapvetően befolyásolja, pontosan meg kell határoznunk az „extrémebb kimenetel” fogalmát. A szerzők a cikkben két kritériumot alkalmaznak, melyeknek egyszerre kell teljesülniük (az így kapott halmaz neve „step down S ”, ennek alakját mutatja a 14. ábra):

1. a hibás tételek száma nem haladja meg a minta hibás tételeinek számát, illetve
2. a hibák összértéke nem haladja meg a minta hibáinak összértékét.

A második lépésben meghatározzuk azon sokasági (p_i) paraméter-együtteseket, melyekre a feltételes valószínűségek összege S halmaz felett legalább akkora, mint az előre rögzített megbízhatósági szint inverze (α). Ezen sokasági paraméter-együttesek halmaza mint konfidencia-halmaz felett maximalizálva $\sum_{i=0}^{100} \frac{i}{100} \cdot p_i$ értékét kapjuk meg a felső határ multinomiális becslését.

Noha ennek a becslésnek sem ismert a valódi megbízhatósági szintje, szimulációk alapján állítható, hogy nagyon közel van a névlegeshez, és emiatt a becslés sokkal pontosabb, mint akár a Stringer-képlet, akár a cella módszer által adott becslés. Mindezen jó tulajdonságai ellenére nem annyira elterjedt, mint az előbbieket, ugyanis a '80-as években még kevés volt a számítástechnikai kapacitás a második lépés konfidencia-halmazának és az azon történő optimalizálásnak a végrehajtásához.

Következtetések

A leírt becslési módszerekkel kapcsolatosan számos kritika vethető fel:

- csupán szimulációkkal bizonyított, hogy valóban legalább a névleges szinten megbízhatóak, ami részben megkérdőjelezi az alkalmazásukat;
- a szimulációk alapján viszont a módszerek túlságosan is konzervatívak, ami jelentősen csökkenti a pontosságukat és így növeli az ellenőrzöttek kockázatát;
- alkalmazásuknak sok olyan előfeltétele (pl. csak túlértékelés lehetséges) van, mellyel az alkalmazók nincsenek tisztában, és így – minden jószándék ellenére – téves eredményekre fognak jutni (különösen veszélyes ez az auditálást támogató szoftverek alkalmazásakor). További gondot okoz az eredmények helyes értelmezése.

A felvetett problémák azonban gyakorlati szempontból kevésbé jelentősek, ugyanis a beszámolók hitelesítésekor maga az értékösszeg-becslés csak a bizonyosság egy részét adja, és egyéb (például elemző) eljárások is megalapozzák a végleges auditori véleményt. A becslések esetleges konzervativizmusát ezen kívül az általánosan elterjedt mintabővítés is kompenzálja, így csökkentve az ellenőrzött szervezetek kockázatát.

A multinomiális becslés és a Stringer-féle képlet elemzése

Ebben a szakaszban a Stringer és a multinomiális becslések implicit elutasítási tartományait hasonlítom össze, és ennek segítségével próbálok közelebb jutni valós megbízhatósági szintjükhöz. A cella módszerről egyrészt egy egész könyv (Leslie–Teitlebaum–Anderson [1980]) szól, így azt nem tartom indokoltnak újraelemezni, másrészt nem is hasonlítható össze a másik kettővel, mivel ehhez a módszerhez saját kiválasztási algoritmus tartozik (cellás kiválasztás), szemben az első kettővel, ahol elvileg az egyszerű véletlen minta is megfelelő / a megfelelő.

A multinomiális felső határ korábban leírt, Fienberg–Neter–Leitch [1977] által javasolt meghatározási módja átfogalmazható a következőképpen:

- Tekintsük a lehetséges $\{p_i\}_{i=0}^{100}$ sokasági hibaarány kombinációk halmazának $(\{p_i : i = 0..100, \sum_{i=0}^{100} p_i = 1\})$ egy olyan partícionálását, amelyet a sokasági hiba átlaga, $\sum_{i=0}^{100} \frac{i}{100} \cdot p_i$ definiál. Az egyes partíciók értelemszerűen reprezentálhatóak egy megfelelő $d \in [0;1]$ számmal, konkrétan a sokasági hiba átlagával. A $d \in [0;1]$ számot, és a partíció egy elemét, tehát a pontosan d hibát tartalmazó sokaságokból álló $\left\{ p_i : i = 0..100, \sum_{i=0}^{100} p_i = 1, \sum_{i=0}^{100} \frac{i}{100} p_i = d \right\}$ halmazt megkülönböztetendő, az utóbbit vastag **d** jelöli.
- Határozzuk meg minden **d** halmazon a $\sum_{\{z_i\} \in S} \frac{n!}{z_0! z_1! \dots z_{100}!} \prod_{i=0}^{100} p_i^{z_i}$ összeg maximumát, azaz a minta S halmazba esésének maximális valószínűségét, ahol $S \subseteq \mathcal{S}$.
- Keressük meg azt a legnagyobb d -t, amire ez a maximum legalább α . Nyilvánvaló, hogy amennyiben ez a maximális d egynél kisebb, a maximum pontosan α -val lesz egyenlő. Ezt az indokolja, hogy a $\sum_{\{z_i\} \in S} \frac{n!}{z_0! z_1! \dots z_{100}!} \prod_{i=0}^{100} p_i^{z_i}$ összeg folytonos a $\{p_i\}_{i=0}^{100}$ halmaz felett.

Az átfogalmazásból már világosabban látszik a multinomiális felső korlát logikája: amennyiben a sokasági hiba várható értéke meghaladja d -t, legfeljebb α valószínűsége van annak, hogy a minta az S halmazba essen, azaz még extrémebb mintát kapjunk. Itt kell

megjegyezni, hogy az alsó határ meghatározása analóg módon történik, és pedig megfelelően definiált S halmaz mellett a legkisebb ilyen d kiválasztásával.

A II. részben definiáltuk az intervallumbecslést megbízhatóságát. Annak a valószínűségnek a minimumát értettük ez alatt a nem-bayesi esetben, hogy a becslt intervallum tartalmazza a keresett sokasági jellemzőt. Figyelembe véve az intervallumbecslés és a hipotézisvizsgálat összefüggését, adott felső határ becslési eljárásnál minden egyes mintához hozzárendelhető azon minták halmaza, amelyekhez kisebb becslt felső határ tartozik: ez a halmaz jelenti a kapcsolódó hipotézisvizsgálat elutasítási tartományát, aminek a mértéke nem haladhatja α -t, a megbízhatósági szint komplementerét.

Az S halmazt a multinomiális felső határ kontextusában az „extrémebb kimenetelű” minták alkotják. Ha a lehetséges minták tere teljesen rendezett, és az S halmazt ennek a rendezésnek megfelelően definiáljuk minden egyes mintánál, akkor biztosak lehetünk abban, hogy a becslésünk megbízhatósága megfelelő. Ennek az az oka, hogy egyrészt az S halmazok a teljes rendezésből való származtatás miatt monoton osztályt alkotnak, tehát mindig fennáll valamilyen tartalmazási reláció bármely két halmaz között, másrészt pedig $S_1 \subseteq S_2$ minden

$$\{p_i\}_{i=0}^{100} \text{ vektorra maga után vonja } \sum_{\{z_i\} \in S_1} \frac{n!}{z_0! z_1! \dots z_{100}!} \prod_{i=0}^{100} p_i^{z_i} \leq \sum_{\{z_i\} \in S_2} \frac{n!}{z_0! z_1! \dots z_{100}!} \prod_{i=0}^{100} p_i^{z_i}$$

teljesülését is, így az elutasítási tartományok mindig egybeesnek a megfelelő S halmazokkal.

Amennyiben az S halmazokat nem teljes rendezésből származtatjuk, akkor már korántsem biztos, hogy a becslés megbízhatósága tényleg annyi, mint amit állítunk róla. Vegyük észre, hogy a step down S definíciójában nem kapunk teljes rendezést a minták között. Ilyenkor:

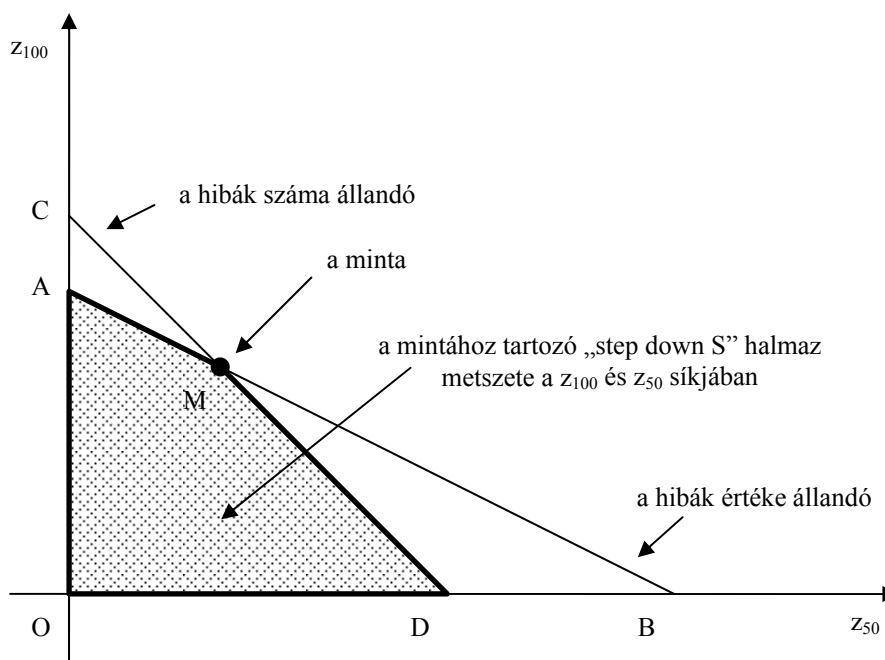
- Biztosan van legalább két olyan mintánk, m_1 és m_2 , amelyek az általuk definiált S_1 és S_2 halmazok segítségével közvetlenül nem összehasonlíthatók: egyik minta sem „extrémebb” a másiknál, tehát nem áll fenn tartalmazási reláció egyik irányban sem az S_1 és S_2 halmazok között. A step down S esetében például nem tudjuk közvetlenül összehasonlítani, hogy 1 db 100%-os + 1 db 80%-os, vagy 2 db 50%-os + 3 db 10%-os hiba-e a rosszabb.
- Ilyen m_1 és m_2 minták egyikéhez, például az m_1 -hez tartozó becslés biztosan nagyobb vagy egyenlő, mint a másik mintához tartozó becslés: a $d(m_1)$ becslés értékének, mint hipotézisnek az elutasítási tartománya tehát tartalmazza m_2 -t, azaz bővebb S_1 -nél.
- Nem teljesen rendezett mintatér esetén tehát a legtöbb d -re, mint „hipotézisre” az eredeti S halmaznál bővebb a valódi elutasítási tartomány. Az elutasítási tartomány

bővülése pedig azzal járhat, hogy csökken a becslésünk megbízhatósága, ugyanis semmi sem biztosítja, hogy az elutasítási tartományba esés valószínűsége ne haladja meg semelyik d sokasági hibát adó esetben sem a szignifikanciaszintet.

Figyeljük meg, hogy egy nem teljesen rendezett mintát utólag teljesen rendezetté tehető az elutasítási tartományok ismeretében, hiszen azok monoton osztályt alkotnak.

Amikor a 70-es 80-as években kifejlesztették a multinomiális módszert, az egyik legfontosabb szempont volt, hogy a felső határ kiszámolható legyen az akkori technikai színvonalon is. Mivel a számításigényt alapvetően az S halmaz mérete határozza meg, nyilvánvaló, hogy minél szűkebbre vesszük az adott mintához tartozó S halmazt, annál gyorsabban kapunk eredményt. Ha csökken az adott mintánál extrémebb kimeneteket tartalmazó S mérete,

csökken $\sum_{\{z_i\} \in S} \frac{n!}{z_0! z_1! \dots z_{100}!} \prod_{i=0}^{100} p_i^{z_i}$ értéke is, tehát egyes mintákra csökkenhet a becsült felső intervallumhatár értéke. Ez azonban nem feltétlenül jelenti azt, hogy a felső határ becslés megbízhatósága csökkenni fog.



14. ábra: Az extrémebb kimenetek step down S halmazának egy metszete a 100%-os és 50%-os szennyezettségek síkjában

A szerzők az eredeti cikkben (Fienberg–Neter–Leitch [1977]) először ugyan teljes rendezéssel (konkrétan a mintabeli hibák összértékével) definiálják az S halmazt, de a konkrét módszerre

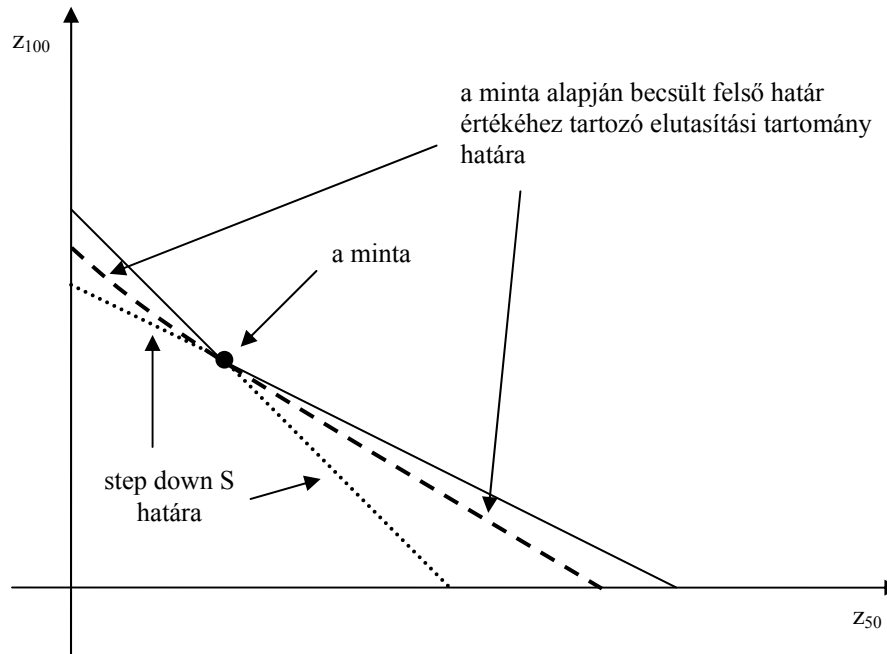
térve – számítási megfontolásokból – átállnak a step down S használatára. A 14. ábra mutatja a step down S halmaz egy, az 50%-os és 100%-os szennyezettségek által kifeszített síkkal párhuzamos metszetét (az egyszerűség kedvéért az ábra figyelmen kívül hagyva a probléma diszkrét jellegét, és folytonos vonalakat használ). A cikket követően tanulmányok, disszertációk sora foglalkozott a módszerrel, kis túlzással megalakult a „Georgia-i iskola”, a University of Georgia-n összegyűjtve a kor monetary unit samplinggal foglalkozó szakembereit. Többen, például Plante [1980] kísérletet tesznek a számítási eljárás algoritmikus javítására, és az alsó határ becsléshez való adaptálására, de (a felső határ esetében) érdemben nem térnek el az eredeti tanulmányban javasolt halmaztól.

Ha – az eredeti gondolatnak megfelelően – *a mintabeli hibák összértékével definiálnánk* az S halmazt (ABO háromszög az ábrán), akkor abban az extrém esetben, amikor a mintában csak 1%-os hibák vannak, vagy minden tétel 0%-nál több hibát tartalmaz, a step down S halmaz (AMDO négyszög) megegyezik a mintabeli hibák összértékén alapuló halmazzal, így ebben az esetben a két becslés is megegyezik. Minden más esetben a step down S halmaz részhalmaza a mintabeli hibák összértékével definiált halmaznak (lásd 14. ábra). Ez azt jelenti, hogy a mintabeli hibák összértékéből származtatott becslés nagyobb vagy egyenlő a step down S halmazokkal való becslésnél, így elutasítási tartománya részhalmaza lesz a step down S halmazokkal definiált multinomiális becslés elutasítási tartományának.

Ha a mintabeli hibák számával definiálnánk az S halmazt (CDO háromszög), a kapott becsléseknek minden minta esetén meg kell egyeznie az egyszerű sokasági arány becslést alapul vevő, tehát mindenhol 100%-os szennyezettséget feltételező (4.5) becsléssel. Tudjuk, hogy konstrukciójából adódóan a step down S halmaz megegyezik ezzel a halmazzal mindazon esetekben, amikor a mintában csak 100%-os szennyezettség van, amiből következik, hogy a multinomiális felső határ is megegyezik ilyenkor a (4.5) becsléssel. A mintabeli hibák számával definiált becslés elutasítási tartományai – hasonlóan az előbbihez – részhalmazai a step down S halmazokkal definiált multinomiális becslés elutasítási tartományának.

A leírt gondolatmenet alátámasztja, hogy a multinomiális felső határ step down S halmazokkal való definiálásakor kapott becslések esetén a becsült értékhez, mint hipotézishez tartozó elutasítási tartományának határa a 14. ábra OCMB négyszög által határolt területből nem léphet ki, de mindenképpen tartalmazza az M mintához tartozó step down S halmazt. Ezt mutatja a 15. ábra is, ahol pontozott vonal jelöli a step down S halmaz határát, és szaggatott vonal jelöli az elutasítási tartomány határát.

Kérdés, hogy ezeknek a megfigyeléseknek a segítségével tudunk-e explicit módon olyan S halmazt definiálni, amelyik a step down S -hez hozzárendelhető elutasítási tartománnyal megegyezik, vagy nagyon hasonló hozzá.



15. ábra: A step down S halmazokkal definiált multinomiális becslés elutasítási tartománya

A pontos egyezés igazolása annyit jelentene, hogy meghatároztuk a step down S -en alapuló multinomiális becslés elutasítási tartományait, és ezáltal a megbízhatósági szintjét is, ezt azonban eddig senkinek sem sikerült általánosan meghatározni. Könnyebb feladat tehát az S halmazok olyan definiálása, amely egyrészt ekvivalencia-osztályokra osztja a mintateret, másrészt nem növeli extrém módon a szükséges számítási igényt. Az ilyen S halmazok monoton osztályt alkotnak, és így nem lesz kétséges a becslés megbízhatósága sem.

Véleményem szerint a legegyszerűbb ilyen definíció a step down S két feltételének lineáris kombinációjával kapható. Az egyszerűség kedvéért megismétlem (ez alkalommal formálisan) a step down S definíciójában szereplő két feltételt – ahogy korábban, $\{z_i\}_{i=0}^{100}$ most is a mintater egy tetszőleges elemét jelöli:

- $\sum_{i=1}^{100} z_i = n - z_0 \leq c$, ahol „ c ” a mintában lévő hibás tételek száma, és
- $\sum_{i=0}^{100} \frac{i}{100} z_i \leq v$, ahol „ v ” a mintában lévő hibák (szennyezettségek) összértéke.

Ezek lineáris kombinációja a következő feltételt adja:

$$\sum_{i=1}^{100} \left(\frac{\lambda \cdot 100 + (1-\lambda) \cdot i}{100} \right) z_i \leq \lambda \cdot c + (1-\lambda) \cdot v \quad (4.11)$$

Ez a feltétel a két szélső esetben megegyezik a sokasági arány becsléssel, illetve a hibák összértékén alapuló becsléssel. Az átmeneti esetekre a felső határ numerikus meghatározását továbbra is segíthetik a Plante doktori értekezésében (Plante [1980]) leírt algoritmikus optimalizálások, ugyanis az ott leírt gondolatok felhasználhatók egy ilyen alakú halmaznál is.

Alternatív, még általánosabb megoldás lehetne $\{z_i\}_{i=0}^{100}$ minta valamilyen CES függvényének⁷³ alkalmazása (ez természetesen tartalmazza speciális esetként a (4.11) formulát is), figyelembe véve azt, hogy a függvény által definiált S halmazok minden mintánál tartalmazzák a step down S halmazt. (Ez utóbbi például a parciális deriváltakra tett pótlólagos feltevésekkel egyszerűen garantálható). A CES függvények használata azonban, noha elméletben vonzó lehet, a kiszámítás bonyolultsága, a szükséges processzoridő miatt a gyakorlat szempontjából nem valós alternatíva.

A Stringer féle felső határ megbízhatóságának elemzése szintén elvégezhető az előbb leírt keretben akkor, ha a multinomiálishoz hasonlóan csak 0, 1, 2, ..., 100 százalékos szennyezettségeket engedünk meg. Ez interpretálható például úgy, hogy a valós szennyezettséget felfelé kerekítettük. Az elemzés célja közelebb jutni a Stringer féle felső határ becslés által generált elutasítási tartományok meghatározásához, hasonlóan a korábban leírt multinomiális esethez.

Amennyiben a mintában csak 100%-os hiba van, akkor a Stringer becslés (akárcsak a multinomiális) megegyezik a sokasági arányon alapuló (4.5) becsléssel. Ez azt jelenti, hogy tetszőleges sokasági hiba-hipotézishez tartozó elutasítási tartomány z_{100} tengelyen vett tengelymetszete megegyezik a Stringer, a multinomiális és a (4.5) becslésnél is. Mivel a Stringer-féle felső határ értéke sohasem haladja meg a (4.5) becsléssel kapott értéket, ezért azt is tudjuk, hogy a Stringer-féle felső határ elutasítási tartománya tartalmazza a (4.5) becslés által implikált elutasítási tartományt.

⁷³ Az általános CES (constant elasticity of substitution) függvényalak használata nagyon elterjedt a közgazdasági elemzésekben, bővebben lásd pl.: Zalai [2000], 277. oldal

A 15. ábra a multinomiális becslés 100 dimenzióban értelmezett (a 0 százalékos szennyezettség nem alkot külön dimenziót!) elutasítási tartományának egy kétdimenziós metszete volt. Hasonló metszetet a (4.7) képletből kiindulva úgy kaphatunk, ha megadjuk a 1...100 százalékos szennyezettségek, mint argumentumok terében a formula által definiált szintvonalakat, majd ezeknek vesszük egy metszetét. Sajnos ez több okból sem tökéletes megoldás, de a legfőbb ok, hogy a probléma diszkrét, és emiatt kénytelenek vagyunk folytonossá tenni a feladatot.

A könnyebb érthetőség kedvéért tekintsük újra a (4.7) képletet, de most kicsit általánosítva:

$$\int_0^n p'(x) t(x) dx \quad (4.12)$$

Az egyszerűség kedvéért elhagytam a minta elemszámára utaló „n” és a megbízhatósági szintre utaló „1- α ” jelöléseket, illetve a konstans „Y” könyv szerinti értékkel való szorzást, és a konstans eltolást is. Ezek a módosítások nem befolyásolják a szintvonalak elhelyezkedését, és a pár bekezdéssel korábban leírtak alapján a csak 100 százalékos szennyezettségű mintáknál a szintvonal értéket amúgy is meg tudjuk adni (de nincs rá szükség).

A $p(x)$ függvény a binomiális felső határ becslés folytonossá tett változatát jelöli, tehát x , a talált hibák száma nem csak egész értéket vehet fel. Kérdés, hogyan lehet $p(x)$ -et meghatározni. Erre az egyik lehetséges megoldást a béta-függvény jelenti, aminek határozatlan integrálként felírt, normált alakja egész argumentumoknál megadja a binomiális eloszlás eloszlásfüggvényét, a következőképpen:

$$F(x | n, p) = 1 - \frac{B_p(\lfloor x \rfloor + 1, n - \lfloor x \rfloor)}{B(\lfloor x \rfloor + 1, n - \lfloor x \rfloor)} = 1 - \frac{\int_0^p t^{\lfloor x \rfloor} (1-t)^{n-\lfloor x \rfloor-1} dt}{\int_0^1 t^{\lfloor x \rfloor} (1-t)^{n-\lfloor x \rfloor-1} dt} \quad (4.13)$$

A $p(x)$ függvény deriváltjának meghatározása az implicit-függvény tétel alapján történhet akkor, ha a tétel alkalmazható. A tétel alkalmazhatóságának feltétele egy tetszőleges (p, x) belső pontban egyebek mellett az, hogy a függvény ott folytonosan deriválható legyen: ez azonban – a képletben szereplő lefele kerekítések miatt – nem garantálható.

Ennek ellenére tegyük fel, hogy sikerült valahogyan (nem feltétlenül a béta-függvényen keresztül) létrehozni egy folytonos $p(x)$ függvényt, és annak deriváltját. Az eredeti, diszkrét $p_{1-\alpha}(i)$ függvény tulajdonságai miatt erre a $p(x)$ függvényre $p(0) = 1 - \sqrt[n]{\alpha}$, $p(n) = 1$, továbbá a függvény monoton növekvő (amíg nem éri el az 1 értéket, addig szigorúan is) és feltételezhető, hogy konkáv is; deriváltja ennek megfelelően nemnegatív, monoton csökkenő függvény lesz.

A $t(x)$ függvény az x -edik legnagyobb szennyezettség értékét mutatja, emiatt monoton csökkenő, lépcsős függvény. Ebben a kontextusban értékei a megengedett 0, 1, 2, ..., 100 százalékos szennyezettségek, más értéket nem vehet fel. Itt következik egy újabb probléma, éspedig az, hogy az elutasítási tartományt a lehetséges $t(x)$ függvények terében kell ábrázolni, tehát a jelen esetben az 1...100 százalékos szennyezettségek mintabeli gyakorisága segítségével, az általuk kifeszített 100 dimenziós vektortérben.

A szennyezettségek és a függvény értékei közötti az alábbi összefüggés áll fenn. A $\{z_i\}_{i=0}^{100}$ minta 101 dimenziós vektora azt fejezi ki, hogy az i % szennyezettséget hány mintaelemnél figyelhettük meg. Mivel a $t(x)$ függvény a szennyezettségek csökkenő sorrendjét fejezi ki, ezért értéke 0 és z_{100} között 1, z_{100} és $(z_{100} + z_{99})$ között 0,99, és így tovább, egészen az $(n - z_0)$ és n közötti szakaszig, ahol 0, hiszen a feltevés szerint csak 0, 1, 2, ..., 100 százalékos szennyezettségek megengedettek.

Felhasználva mindezeket, a (4.12) formula az alábbi módon írható át:

$$\sum_{i=0}^{100} \left((100-i) \int_{\sum_{j=100-i+1}^{100} z_j}^{\sum_{j=100-i}^{100} z_j} p'(x) dx \right), \text{ ahol } z_{101} = 0. \quad (4.14)$$

Amennyiben meg szeretnénk határozni konkrét elutasítási tartomány határfelületének egy tetszőleges síkmetszetét, akkor ezt a (4.14) formulából az implicit függvény tétel segítségével (vagy numerikusan) meg lehet tenni tetszőleges z_j z_k szennyezettségekre. Sajnos ez az eredmény nem alkalmas arra, hogy belőle továbblépve meghatározzuk általános érvénnyel a Stringer becslés elutasítási tartományát. Ennek a szakasznak azonban nem is az volt célja, hogy pontosan meghatározza a becslés elutasítási tartományát, hanem csak az, hogy felvázoljon egy lehetséges új megközelítést az ilyen típusú problémák későbbi megoldásához.

Láttuk tehát, hogy mind a multinomiális, mind a Stringer féle felső határ becslések elutasítási tartományai tartalmazzák a csupán a hibák mintabeli darabszámát felhasználó (4.5) becslés elutasítási tartományát, ami azt mutatja, hogy az általuk becsült felső határ-érték kevésbé konzervatív. Arra azonban nem sikerült ezúttal választ találni, hogy vajon ezek a becslések tényleg minden esetben megbízhatóak-e.

A jövőre nézve, az elmélet szempontjából véleményem szerint nem volt hiábavaló az itt leírt megközelítés. Egy lehetséges utat jelent a megbízhatóság bizonyítására, például úgy, hogy egyrésztől találunk olyan monoton osztály halmazrendszert, ami tartalmazza a multinomiális

felső határ megfelelő elutasítási tartományait, másrésztől be tudjuk bizonyítani, hogy a multinomiális felső határ elutasítási tartományai tartalmazzák a Stringer-féle megfelelő tartományokat. Az első ugyanis azt jelenti, hogy a step down S-re alapozott multinomiális felső határ megbízható, míg a második az elsővel együtt a Stringer becslés megbízhatóságát jelenti.

A tényleges gyakorlat szempontjából azonban, ahogy az előző szakasz végén is leírtam, az itt említetteknek kicsi a jelentősége. A programcsomagok továbbra is a Stringer, vagy a cellás becslést fogják tartalmazni, legalábbis addig mindenképpen, amíg a számítástechnika fejlődése lehetővé nem teszi multinomiális felső határ személyi számítógépen való kiszámítását. Az audit szempontjából pedig teljesen elégséges az az ésszerű bizonyosság, amit a becslések megbízhatóságáról elvégzett szimulációk mutatnak, hiszen a valóságban súlyosabb tényezők is vannak a feltárási kockázatnak, mint pár százalékpont tévedés a mintavételi kockázat meghatározásában.

Összegzés

Az értekezés során igyekeztem áttekintést nyújtani a jelenleg használt ellenőrzési technikákról. Ezekkel kapcsolatosan ugyan számos kritika vethető fel (ezek egy részét több helyen fel is vetettem), de kétségtelen, hogy az ellenőrzési rendszerek többsége, ha nem is tökéletesen, de betölti a funkcióját. Más kérdés, hogy ezt a funkciót be lehetne-e tölteni hasznosabban, nagyobb bizonyosságot nyújtva, és esetleg mindeközben kevesebb erőforrást felhasználva. Az értekezés éppen az utóbbi kérdésre próbált választ adni, éspedig oly módon, hogy az ellenőrzési módszerek statisztikai-valószínűségelméleti hátterét megvizsgálja.

A jelenleg használt módszerek nem tökéletesek, és véleményem szerint soha nem is lesznek azok. Lehetőség van azonban apróbb, ellenőrzéstechnikai és nem utolsósorban erőforrás-felhasználási szempontból fontos optimalizálásokra, amint azt a III. rész végén bemutatott, az Állami Számvevőszék átfogó ellenőrzéseinél most bevezetett szekvenciális jellegű aránybecslési módszer kezdeti tapasztalatai is mutatják. Ez az újítás nem születhetett volna meg az aránybecslési eljárások alapos tanulmányozása, az értekezésben is leírt elemzése nélkül. Mivel a Sarbanes Oxley Act különös aktualitást ad a kontrollok tesztelésének, ami akár USA, akár világgazdasági szinten egy összességében meglehetősen költséges feladat, a kontrolltesztelés hatékonyságának növelése közvetlen, és pénzben is kifejezhető előnyökkel jár. Noha a szekvenciális módszerek bevezetésére korábban is történtek kísérletek, a bemutatott módszer könnyebb interpretálhatósága (főleg a negatív megállási pontok esetében) és a 2007. eleje óta felhalmozódott gyakorlati tapasztalatok könnyebben elfogadhatóvá teszik ezt a módszert bármely ellenőrzési szakember számára, mint a Wald féle (amúgy optimális) szekvenciális tesztet. Jelenleg előkészületben van egy angol nyelvű publikáció, amely a tesztelési módszer leírását és az első év alkalmazási tapasztalatait tartalmazza.

A kockázatelemzés olyan terület, ahol igazán sok hozzáadott értéket tud jelenteni egy jól megválasztott statisztikai módszer: leginkább a piaci kockázat kezelésére kifejlesztett banki rendszereknél tapasztalhatjuk ezt. A kockázatelemzés az ellenőrzési munkában is fontos: ha jól alkalmazzák, nagyon sok erőforrás megtakarítható általa. Az esetleges jövőbeni fejlesztések nem állhatnak meg ott, hogy objektíven, a korábbi ellenőrzési tapasztalatokból határozzuk meg a kockázati súlyokat (noha ennek elérése is jelentős előrelépést jelentene a

jelenlegi helyzethez képest), hanem a konkrét alkalmazási terület ismeretében új kockázati modelleket, új függvényformákat kell kidolgozni.

Bízom benne, hogy kutatásnak számos olyan ellenőrzésszakmai kérdésre is sikerült rávilágítania, amelyek főleg elméleti szempontból tekinthetők fontosnak, legalábbis jelenleg.

Az első ilyen eredmény szintén a kockázatelemzéshez kapcsolódik. Az I. és a II. részben is szó volt róla, hogy az eredendő- és a kontrollkockázatok belső struktúrája igen változatos lehet, és túlzottan leegyszerűsítő ezzel kapcsolatban az összes használt modell, ami rossz esetben a kockázatok súlyos alábecslésével is járhat, főleg a vezetők által elkövetett csalások esetén (management override of controls). A jövőre nézve érdekes kutatási irány lehetne annak feltárása, hogy a kockázati struktúra általánosabb modellje hogyan illeszthető be, beilleszthető-e egyáltalán a gyakorlati alkalmazásokba.

A Benford-törvény és a karakterisztikus függvény kapcsolatát bemutató tétel, illetve az ebből következő új elméleti magyarázat a tétel gyakorlati érvényesülésére egy újabb megerősítést tartalmaz arra nézve, hogy a való életből származó eloszlásoktól ténylegesen elvárható a Benford törvénynek megfelelő viselkedés: ez is azt támasztja tehát alá, hogy a számjegyelemzés egy legitim kockázatelemzési módszer. Sajnálatos módon „éles” kísérleteket nem tudtam végezni a számjegyelemzéssel, ugyanis az ÁSZ jelenlegi ellenőrzési stratégiájába ennek a módszernek az alkalmazása nem fér bele.

Az értékösszeg-becslési rész eredményeinek jelenleg csak elméleti jelentősége van, a multinomiális felső határ megbízhatóságáról szóló eszmefuttatás a számítástechnika fejlődése útján, az egzakt becslések tömeges elterjedésével nyerheti el az értelmét.

Reményeim szerint ez a munka nem volt öncélú, és az itt leírtak hasznosulni fognak a jövőben oktatási segédanyagként, esetleg új módszertani útmutatók kidolgozásakor, a régiek frissítésekor.

Különösen fontosak az ilyen típusú kutatások az olyan közpénzből működő ellenőrző szervezeteknél, mint amilyen az Állami Számvevőszék, ahol a feladatok folyamatosan növekvő mennyiségével az ellenőrzések hatékonyságának fokozása által lehet lépést tartani. A hatékonyság fokozásának útja pedig nem más, mint egyre kifinomultabb statisztikai módszerek alkalmazása, melyek segítenek ugyanazt a magas szintű bizonyosságot kevesebb ráfordítással elérni. És azt, hogy ez mennyire nem egy lényegtelen apróság, az is mutatja, hogy egy ellenőri nap költsége a 2007. évben körülbelül 110 000 Ft.

A dolgozat végén szeretném ismételtén felhívni a figyelmet azokra a veszélyekre, amelyek együtt járnak a módszertani fejlesztésekkel. Az új ellenőrzési eljárások alkalmazása, különösen a szoftverek segítségével történő auditálás, jelentős kockázatokat hordoz az ellenőrző szerveknél. Fontos, hogy az ellenőrök szabatosan használják a szakkifejezéseket, és értsék is azt, amit csinálnak: ez különösen igaz az értékösszeg-becslés módszertanilag elég kényes területére. Az auditálást támogató számítógépes programok ugyan óriási segítséget jelentenek, de egyben óriási veszélyt is hordoznak magukban: mechanikusan, az algoritmus ismerete nélkül alkalmazva egy szoftvert könnyen juthatunk téves, vagy megalapozatlan következtetésekre. Ezt a kockázatot csak úgy lehet kiküszöbölni, ha minden újítás bevezetésekor átfogó képzésben részesítjük az alkalmazókat, és a kezdeti időszakban nagyobb hangsúlyt fektetünk a minőségi felülvizsgálatra, az ellenőri tevékenység belső kontrolljára.

Noha túlmutat az értekezés keretein, de az ellenőrzéssel a gyakorlat mellett elméletben is foglalkozók talán egyet fognak érteni azzal a – hazai szakirodalomban már többszörösen leírt – megállapítással, hogy jelenleg a hazai ellenőrzés-szakma egyik súlyos problémája a terminológiai zavar, ami sajnos komoly akadályt képez annak, hogy – az előző bekezdésben leírtakkal összhangban – „az ellenőrök szabatosan használják a szakkifejezéseket”. Ennek a zavarnak az oka messzire vezet, ide tartozik például az is, hogy az évszázados német alapokat egyszerre érték angolszász és francia hatások a rendszerváltás után, attól függően, hogy melyik intézménynek melyik országgal volt testvérintézményi (ún. twinning) kapcsolata.

Nem lehet szó nélkül elmenni olyan jelenségek mellett, hogy a Magyar Könyvvizsgálói Kamara által gondozott hazai könyvvizsgálati szabványok (amelyek alapjában véve az IFAC szabványok fordításai) az angol szakkifejezéseket félreértésre okot adó módon fordítják, így például ugyanaz a szakkifejezés teljesen mást jelent az Állami Számvevőszék, és mást a Kamara szóhasználatában.⁷⁴ A terminológiai zavar azonban nem áll meg ezen a szinten, sajnos – és erre már a bevezetőben is utaltam – egészen a jogszabályokig, magáig az Alkotmányig is elér, ami azzal jár, hogy a végrehajtandó feladatot nem a megrendelő, hanem a végrehajtást végző fogja meghatározni. A jelen értekezés és az ehhez hasonló munkák ugyan segíthetnek kis lépésekkel előre lépni a kereteken belül, de az alapvető problémákat sajnos nem tudják megoldani. Igaz, nem is feladatuk.

⁷⁴ Erre a szakmai körökben leginkább közismert példa az „internal control”

Irodalomjegyzék

- ADHIKARI, A. K. – SARKAR, B. P. [1968] Distribution of most significant digit in certain functions whose arguments are random variables. *Sankhya Ser. B.* 30. évf. 47-58. old.
- ADHIKARI, A. K. [1969] Some results on the distribution of the most significant digit. *Sankhya Ser. B.* 31. évf. 413-420. old.
- ALLAART, P. C. [1997]: An invariant-sum characterization of Benford's law. *Journal of Applied Probability.* 34. évf. 1. sz. 288–291. old.
- ARENS, A. A. – LOEBBECKE, J. K. [1997]: Auditing: An integrated approach. Prentice-Hall. London.
- ASHBAUGH-SKAIFE, H. – COLLINS, D. W. – KINNEY, W. R. JR. – LAFOND, R. [2007]: The Effect of SOX Internal Control Deficiencies on Firm Risk and Cost of Equity. SSRN. Munkaanyag
- ÁSZ MÓDSZERTANI KIADVÁNYOK – LÉVAI J. (szerk.) [2004]: A számvevőszéki ellenőrzés szakmai szabályai. Állami Számvevőszék. Budapest
- BENFORD, F. [1938]: The law of anomalous numbers. *Proceedings of the American Philosophical Society.* 78. évf. 4. sz. 551–572. old.
- BERGER, A. – BUNIMOVICH, L. A. – HILL, T. P. [2005]: One-dimensional dynamical systems and Benford's law. *Transactions of the American Mathematical Society.* 357. évf. 1. sz. 197–219. old.
- BERGER, A. [2004]: Dynamics and digits: on the ubiquity of Benford's law. Vienna University of Technology. Munkaanyag.
- BICKEL, P. J. [1992]: Inference and auditing: The Stringer bound. *International Statistical Review.* 60. évf. 2. sz. 197–209. old.
- CARSLAW, C. [1988]: Anomalies in income numbers: Evidence of goal oriented behavior. *The Accounting Review.* 63. évf. 2. sz. 321–327. old.
- CASEWARE IDEA RESEARCH DEPARTMENT [2003]: Monetary unit sampling technical specification. <http://www.caseware-idea.com>.
- CASEWARE IDEA RESEARCH DEPARTMENT [2003]: White papers on attribute sampling technical specification. <http://www.caseware-idea.com>.
- COCHRAN, W. G. [1977]: Sampling Techniques. 3. kiadás. Wiley. New York.
- CUSHING, B. E. – LOEBBECKE, J. K. [1983] Analytical Approaches to Audit Risk: A Survey and Analysis. *Auditing: A Journal of Practice and Theory.* 3. évf. 1. sz. 23–41. old.

- DAVID, H. A. [1981]: Order statistics. Wiley. New York.
- DE JAGER, N. G. – PAP GY. – VAN ZUIJLEN, M.C.A. [1997]: Facts, phantasies and a new proposal concerning the Stringer bound. Computers and Mathematics with Applications. 33. évf. 10. sz. 37–54. old.
- DENKINGER G. [1990]: Valószínűségszámítás. Tankönyvkiadó. Budapest.
- DHAR, V., LEWIS, B., PETERS, J. [1987]: A Knowledge-Based Model Of Audit Risk. Stern School of Business (New York University) working paper series. IS-87-116
- DIACONIS, P. [1977]: The distribution of leading digits and uniform distribution mod 1. Annals of Probability. 5. évf. 1. sz. 72–81. old.
- DRAKE, P. D. – NIGRINI, M. J. [2000]: Computer assisted analytical procedures using Benford's law. Journal of Accounting Education. 18. évf. 2. sz. 127-146. old.
- DUKE, G. L. [1980]: An empirical investigation of the power of statistical sampling procedures used in auditing under different models of change of error patterns. University of Georgia. Athens. Ph.D. értekezés
- DURTSCHI, C. – HILLISON, W. – PACINI, C. [2004]: The effective use of Benford's law to assist in detecting fraud in accounting data. Journal of Forensic Accounting. 5. évf. 17-34. old.
- FELLER, W. [1971]: An introduction to probability theory and its applications. Wiley. New York.
- FIENBERG, S. E. – NETER, J. – LEITCH, R. A. [1977]: Estimating the total overstatement error in accounting populations. Journal of the American Statistical Association. 72. évf. 295–302. old.
- FLEHINGER, B.-J. [1966]: On the probability that a random integer has initial digit A. American Mathematical Monthly. 73. évf. 10. sz. 1056–1061. old.
- GOODFELLOW, J. L. – LOEBECKE, J. K. – NETER, J. [1974]: Some perspectives on CAV sampling plans I-II. CA Magazine. October, 23–30. old., November, 46–53. old.
- HALDENE, J. B. S. [1945]: On a method of estimating frequencies. Biometrika. 33. évf. 222–225. old.
- HAM, J. – LOSELL, D. – SMIELIAUSKAS, W. [1985]: An empirical study of error characteristics in accounting populations. Accounting Review. 60. évf. 387–406. old.
- HANSEN, M. H. – HURWITZ, W. N. [1943]: On the theory of sampling from finite populations. Annual Mathematical Statistics. 14. évf. 4. sz. 333–362. old.
- HASKINS, M. E. – DIRSMITH, M. W. [1995]: Control and inherent risk assessments in client engagements: An examination of their interdependencies. Journal of Accounting and Public Policy. 14. évf. 1. sz. 63-83. old.

- HILL, T. P. [1988]: Random-number guessing and the first digit phenomenon. *Psychological Reports*. 62. sz. 967–971. old.
- HILL, T. P. [1995a]: Base-invariance implies Benford's law. *Proceedings of the American Mathematical Society*. 123. évf. 3. sz. 887–895. old.
- HILL, T. P. [1995b]: The significant-digit phenomenon. *American Mathematical Monthly*. 102. évf. 4. sz. 322–326. old.
- HILL, T. P. [1996]: A statistical derivation of the significant-digit law. *Statistical Science*. 10. évf. 4. sz. 354–363. old.
- HILL, T. P. [1999]: The difficulty of faking data. *Chance* 26. évf. 1. sz. 8–13. old.
- HILL, T. P.– SCHÜRGER, K. [2005]: Regularity of digits and significant digits of random variables. *Journal of Stochastic Processes and their Applications*. 115. évf. 10. sz. 1723–1743. old.
- HORVITZ, D. G. – THOMPSON, D. J. [1952]: A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*. 47. évf. 12. sz. 663–685. old.
- HUNYADI L. [2001]: Statisztikai következtetésemélet közgazdászoknak. Központi Statisztikai Hivatal. Budapest.
- HUNYADI L.–VITA L. [2004]: Statisztika közgazdászoknak. Központi Statisztikai Hivatal. Budapest.
- INTERNATIONAL FEDERATION OF ACCOUNTANTS (IFAC) [2006]: Handbook of international auditing, assurance, and ethics pronouncements. <http://www.ifac.org>
- JOHNSON, J. R. – LEITCH, R. A. – NETER, J. [1981]: Characteristics of errors in accounts receivable and inventory audits. *Accounting Review*. 56. évf. 270–293. old.
- KAPLAN, R. S. [1973a]: Stochastic model for auditing. *Journal of Accounting Research*. 11. évf. 38–46. old.
- KAPLAN, R. S. [1973b]: Statistical sampling in auditing with auxiliary information estimators. *Journal of Accounting Research*. 11. évf. 238–258. old.
- KATZ, L. [1953]: Confidence intervals for the number showing a certain characteristic in a population when sampling is without replacement. *Journal of the American Statistical Association*. 48. évf. 256–261. old.
- KNUTH, D. [1981]: *The Art of Computer Programming*. Vol. 2. Seminumerical Algorithms. Addison-Wesley. Reading.
- LEHMANN, E. L. [1959]: *Testing statistical hypotheses*. Wiley. New York.
- LESLIE, D. A. – TEITLEBAUM, A. D. – ANDERSON, R. J. [1980]: *Dollar-Unit Sampling-A practical guide for auditors*. Pitman. London.

MAXIM, L. D. – CULLEN, D. E. – COOK, F. X. Jr. [1976]: Optimal Acceptance Sampling Plans for Auditing: "Batched" Stop and Go vs. Conventional Single-Stage Attributes Plans. *The Accounting Review*, 51. évf. 1. sz. 97-109. old.

MEDVEGYEV P. [2002]: Valószínűségszámítás. Aula. Budapest

MILLER, S. J. – NIGRINI, M. J. [2006]: Differences between independent variables and almost Benford behaviour. Arxiv preprint math.PR/0601344 at www.arxiv.org.

MOELLER, R. [2007]: COSO Enterprise Risk Management: Understanding the New Integrated ERM Framework. Wiley. New York.

NETER, J. – KIM, H. S. – GRAHAM, L. E. [1984]: On combining Stringer bounds for independent monetary unit samples from several populations. *Auditing*. 4. évf. 1. sz. 74–88. old.

NETER, J. – LOEBBECKE, J. [1977] On the behavior of statistical estimators when sampling accounting populations. *Journal of the American Statistical Association*. 72. évf. 501–507. old.

NETER, J.–LOEBBECKE, J. [1975]: Behavior of major statistical estimators in sampling accounting populations – An empirical study. AICPA. New York.

NEWCOMB, S. [1881]: Note on the frequency of use of the different digits in natural numbers. *American Journal of Mathematics*. 4. évf. 1. sz. 39–40. old.

NEYMAN, J. [1934]: On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society*. 97. évf. 558–606. old.

NIGRINI, M. J. – MITTERMAIER, L. [1997]: The use of Benford's law as an aid in analytical procedures. *Auditing: A Journal of Practice and Theory*. 16. évf. 2. sz. 52–67. old.

NIGRINI, M. J. [1992]: The detection of income tax evasion through an analysis of digital frequencies. University of Cincinnati. Cincinnati, OH. Ph.D. értekezés

NIGRINI, M. J. [1996]: A taxpayer compliance application of Benford's law. *Journal of the American Taxation Association*. 18. évf. 1. sz. 72–91. old.

NIGRINI, M. J. [2000]: Continuous Auditing. Munkaanyag.

NYIKOS L., BODONYI M., MALATINSZKYNÉ L. I., MÁRKUS G. [2002]: Közpénzek ellenőrzése I-II. Perfekt. Budapest

Panel on Nonstandard Mixtures of Distributions – TAMURA, H. ET AL. [1989]: Statistical models and analysis in auditing. *Statistical Science*. 4. évf. 1. sz. 2–33. old.

PAP GY. – VAN ZUIJLEN, M. C. A. [1995]: The Stringer bound in case of uniform taintings. *Computers Mathematical Applications*. 29. évf. 10. sz. 51–59. old.

PAP GY. – VAN ZUIJLEN, M. C. A. [1996]: On the asymptotic behaviour of the Stringer bound. *Statistica Neerlandica*. 50. évf. 3 sz. 367–389. old.

- PAP GY. – VAN ZUIJLEN, M. C. A. [2000]: Modified Stringer bounds. *Publicationes Mathematicae*. Debrecen. 57. évf. 1–2. sz. 163–183. old.
- PINKHAM, R. S. [1961]: On the distribution of first significant digits. *Annals of Mathematical Statistics*. 32. évf. 4. sz. 1223–1230. old.
- PLANTE, R. [1980]: Algorithmic improvements and extensions for obtaining the multinomial bound on total overstatement and understatement errors. University of Georgia. Athens. Ph.D. értekezés
- PLANTE, R. – NETER, J. – LEITCH, R. A. [1984]: A lower multinomial bound for the total overstatement error in accounting populations. *Management Science*. 30. évf. 1. sz. 37. old
- PRÉKOPA A. [1972]: Valószínűségelmélet. Műszaki Könyvkiadó. Budapest
- POSCH, P. N. [2004]: Ziffernanalyse in der Fälschungsaufspürung. Benford's Gesetz und Steuererklärungen in Theorie und Praxis. Munkaanyag.
- RAIMI, R. A. [1969]: On the distribution of first significant figures. *American Mathematical Monthly*. 76. évf. 4. sz. 342–348. old.
- RAIMI, R. A. [1976]: The first digit problem. *American Mathematical Monthly*. 83. évf. 7. sz. 521–538. old.
- RAIMI, R. A. [1985]: The first digit phenomenon again. *Proceedings of the American Philosophical Society*. 129. évf. 2. sz. 211–219. old.
- ROOT, S. J. [2000] *Beyond Coso: Internal Control to Enhance Corporate Governance*. Wiley. New York.
- SCHRÄPLER, J.-P. – WAGNER, G. G. [2003]: Identification, characteristics and impact of faked interviews in surveys. IZA Discussion paper 969. www.iza.org.
- STRINGER, K. W. [1963]: Practical aspects of statistical sampling in auditing. *Proceedings of Business Economics, Statistics Section*. American Mathematical Association. Washington. Munkaanyag.
- STRINGER, K. W. [1979]: Statistical sampling in auditing. The state of art. *Annual Accounting Review*. 1. sz. 113–127. old.
- THOMAS, J. K. [1989]: Unusual patterns in reported earnings. *The Accounting Review*. 64. évf. 4. sz. 773–787. old.
- VANCE, L. L. [1951]: How much test checking is enough? *The Accounting Review*, 26. évf. 1. sz. 22–30. old.
- VAN HEERDEN, A. [1961]: Steekproeven als Middel van Accountantscontrole. *Maandblad voor Accountancy en Bedrijfshuishoudkunde*. 11. sz. 453. old.
- VARIAN, H. R. [1972]: Benford's law. *The American Statistician*. 26. évf. 3. sz. 65–66. old.

WALD Á. [1945]: Sequential Tests of Statistical Hypotheses. The Annals of Mathematical Statistics, 16. évf. 2. sz. 117-186. old.

WALD Á. [1947]: Sequential Analysis. Wiley. New York.

WALD Á. – WOLFOWITZ, J. [1948]: Optimal Character of the Sequential Probability Ratio Test. The Annals of Mathematical Statistics, 19. évf. 3. sz. 326-339. old.

WALLER, W. S. [1993]: Auditors' Assessments of Inherent and Control Risk in Field Settings. The Accounting Review, 68. évf. 4. sz. 783-803. old.

WRIGHT, T. [1990]: When zero defectives appear in a sample: upper bounds on confidence coefficients of upper bounds. The American Statistician. 44. évf. 40–41. old.

WRIGHT, T. [1991]: Exact confidence bounds when sampling from small finite universes. Springer. New York.

WRIGHT, T. [1992]: A note on sampling to locate rare defectives with strong prior evidence. Biometrika. 79. évf. 685–691. old.

WRIGHT, T. [1997]: A simple algorithm for tighter exact upper confidence bounds with rare attributes in finite universes. Statistics & Probability Letters. 36. évf. 59–67. old

ZALAI E. [2000]: Matematikai közgazdaságtan, KJK-KERSZÖV, Budapest.

Saját publikációk jegyzéke

LOLBERT T. [2003]: Az "objektív" kockázatelemzési módszer. Ellenőrzési Figyelő. 2003/4. sz. 42-45. old.

LOLBERT T. [2004]: A sokasági arány meghatározására irányuló statisztikai eljárások véges sokaság és kis minták esetén. Statisztikai Szemle. 82. évf. 12. sz. 1053–1076. old.

LOLBERT T. [2006a]: A sokasági értékösszeg becslése a könyvvizsgálatban. Statisztikai Szemle. 84. évf. 3. sz. 225–248. old.

LOLBERT T. [2006b]: Digital analysis: Theory and applications in auditing. Statisztikai Szemle. 10. különszám 148–170. old.

LOLBERT T. [2007]: On the non-existence of a general Benford's law. Mathematical Social Sciences. Elfogadva (2007. szeptember 10.)

LOLBERT T. [2008]: Control testing with stopping points: an efficient way to comply with SOx 404. Előkészületben

Függelék

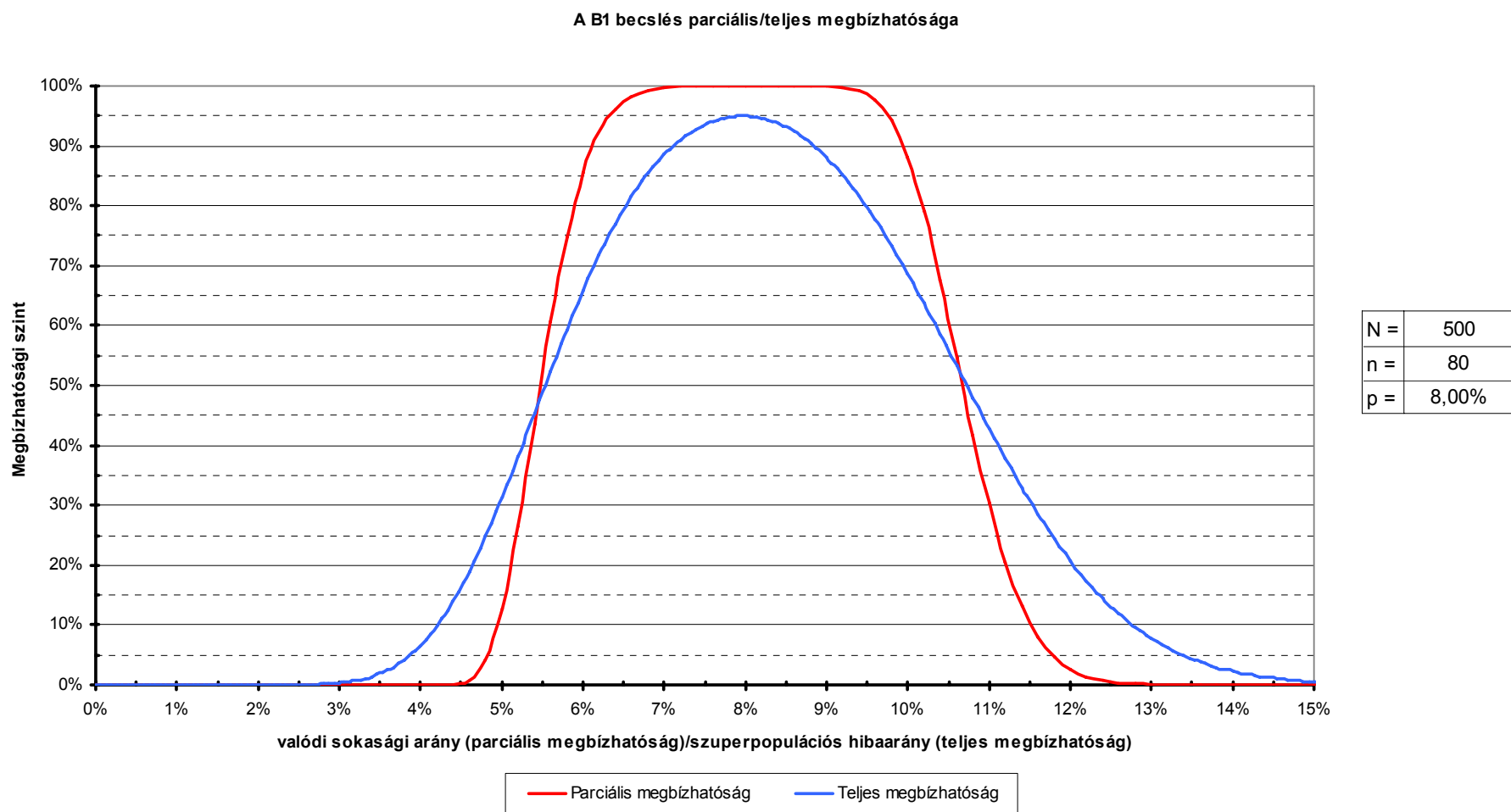
M=

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
0	0,95	0,96875	0,960918	0,951098	0,93878	0,923321	0,903909	0,87952	0,848863	0,810308	0,761794	0,700715	0,623778	0,526814	0,404543	0,25028	0,055548	0	0	0
1	0	0,875	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0,956991	0,893353	0,820218
2	0	0	0,369949	0,820196	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	0	0	0	0	0,066102	0,632869	0,893555	1	1	1	1	1	1	1	1	1	1	1	1	1
4	0	0	0	0	0	0	0	0,261657	0,662465	0,901701	1	1	1	1	1	1	1	1	1	1
5	0	0	0	0	0	0	0	0	0	0	0,248621	0,622772	0,867544	1	1	1	1	1	1	1
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0,147506	0,533021	0,796503	0,992493	1	1	1
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,395849	0,688482	0,907209
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

m=

16. ábra: Az M3 becslés megjelenítése Excelben

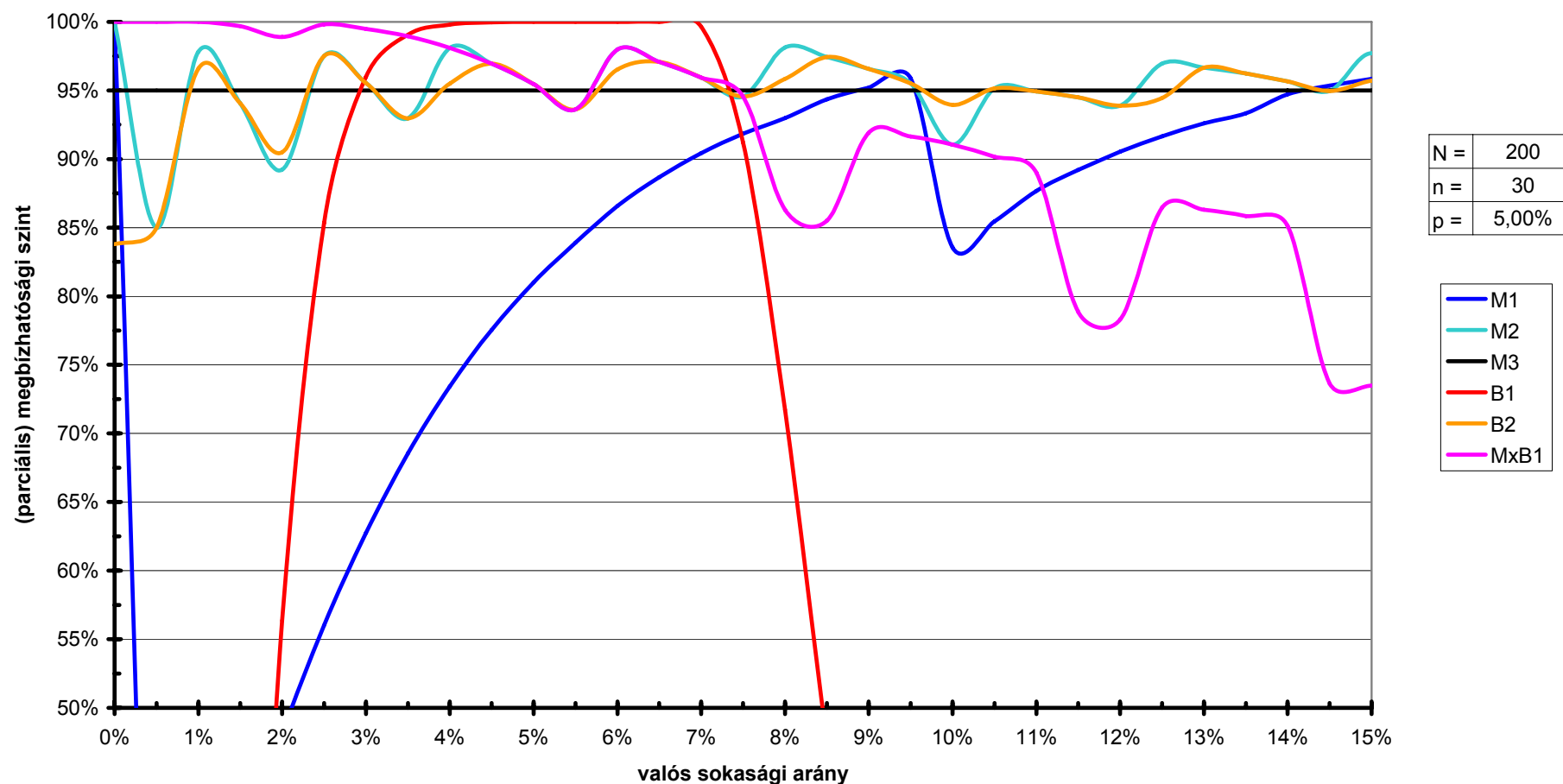
Így nézett ki az elemzésre használt Excel táblában például a randomizált M3 becslés/hipotézisvizsgálat. Becslésként olvasva a cellákba írt számok jelentik a sornak megfelelő minta esetén annak valószínűségét, hogy az oszlopnak megfelelő sokasági hibamennyiség szerepel az intervallumbecslésben. Ugyanez a táblázat olvasható természetesen hipotézisvizsgálatként is; ilyenkor a cellák a hipotézisnek megfelelő oszlopban mutatják a sornak megfelelő minta elfogadási tartományba esésének a valószínűségét. (A világos színű cellák a randomizáltak.)



17. ábra: A B1 becslés parciális és teljes megbízhatósága

A parciális megbízhatósági szint ugyanabból a sokaságból való ismételt mintavételt feltételez, ami a bayesi becslések esetén nem ad értelmezhető eredményt. A teljes megbízhatósági szint ezzel szemben figyelembe veszi azt is, hogy nem csak a mintabeli érték, hanem a sokasági paraméter is valószínűségi változó. Látható, hogy minél messzebb van egymástól a valós és a használt prior, annál kisebb lesz a becslés megbízhatósága.

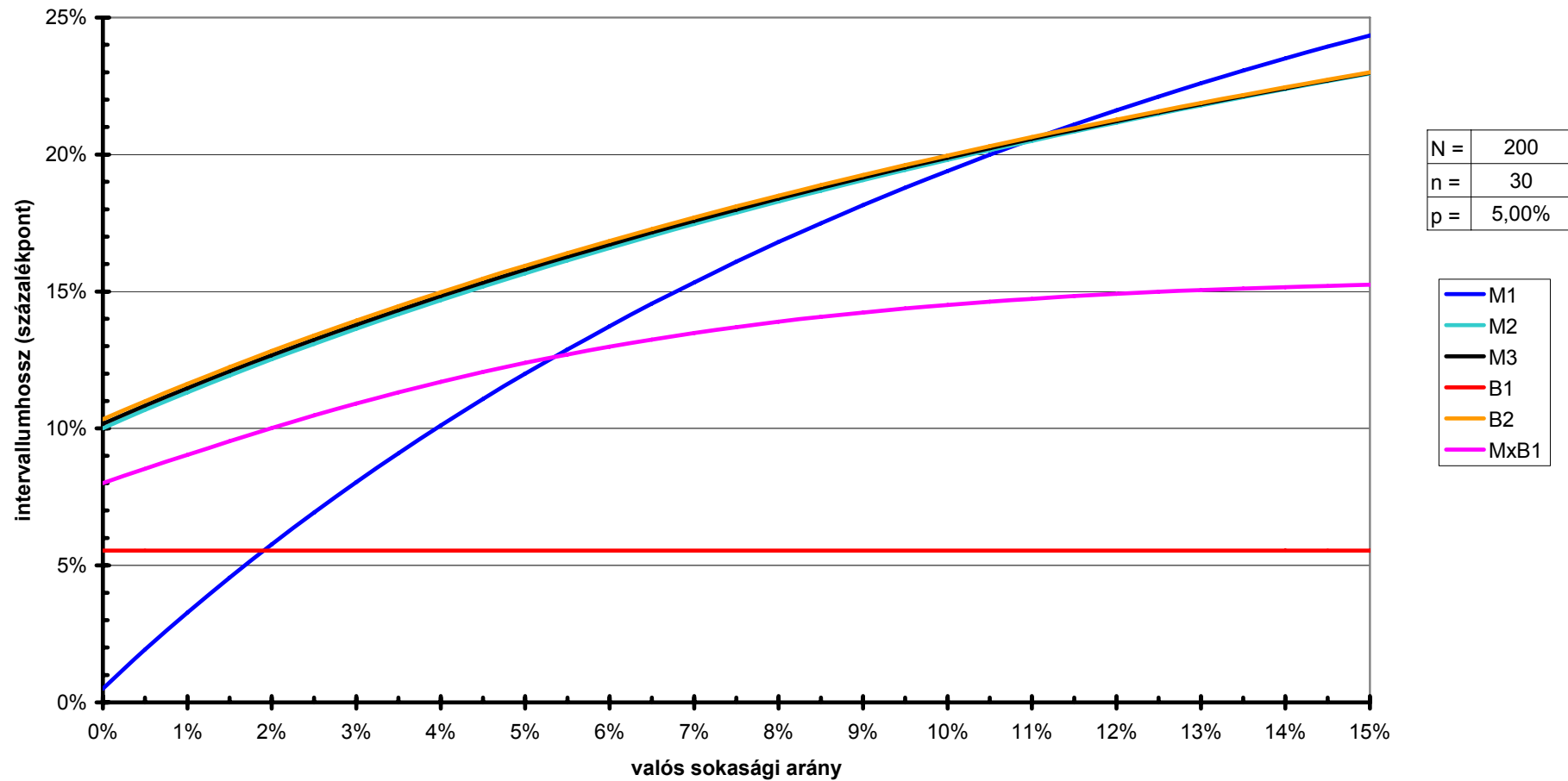
Az intervallumbecslési módszerek (parciális) megbízhatósága



18. ábra: Az elemzés I. outputja a (200/30/5%) forgatókönyv esetén

A parciális megbízhatóság azt fejezi ki, hogy adott sokasági arány mellett nagyon sokszor megismételve a mintavételt, a becslés az esetek hány százalékában tartalmazta a keresett sokasági jellemzőt, tehát a valós sokasági arányt. Torzítatlan becslés esetén ez nem lehetne kevesebb, mint a névleges szint (ami ebben az esetben 95%)

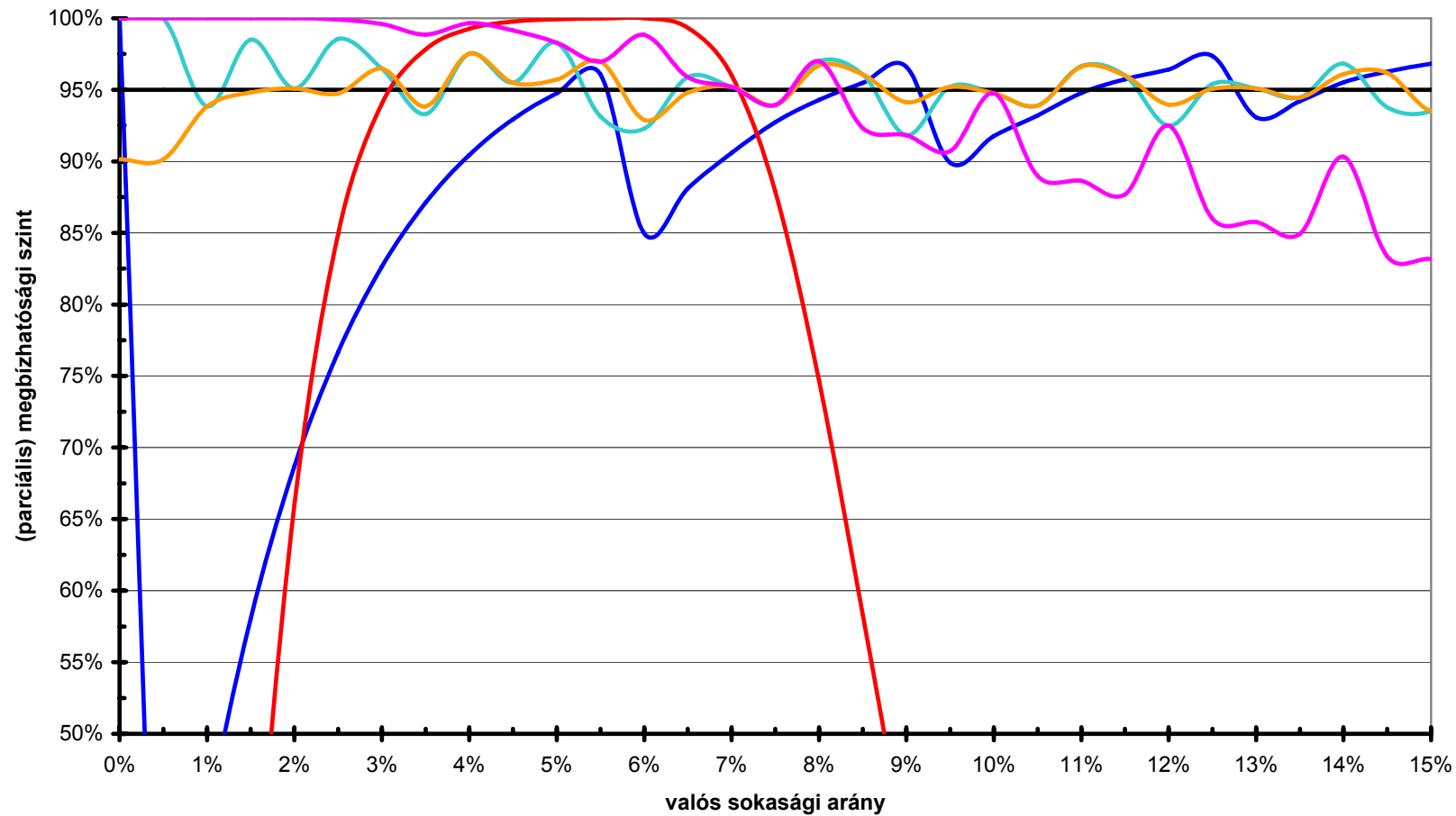
Az intervallumbecslési módszerek várható intervallumhossza



19. ábra: Az elemzés II. outputja a (200/30/5%) forgatókönyv esetén

A várható intervallumhossz a becsült intervallum hosszának várható értékét fejezi ki adott sokasági paraméter mellett. Hosszabb intervallum nem feltétlenül jelent nagyobb megbízhatóságot, ami az előző ábrával összehasonlítva is látható.

Az intervallumbecslési módszerek (parciális) megbízhatósága

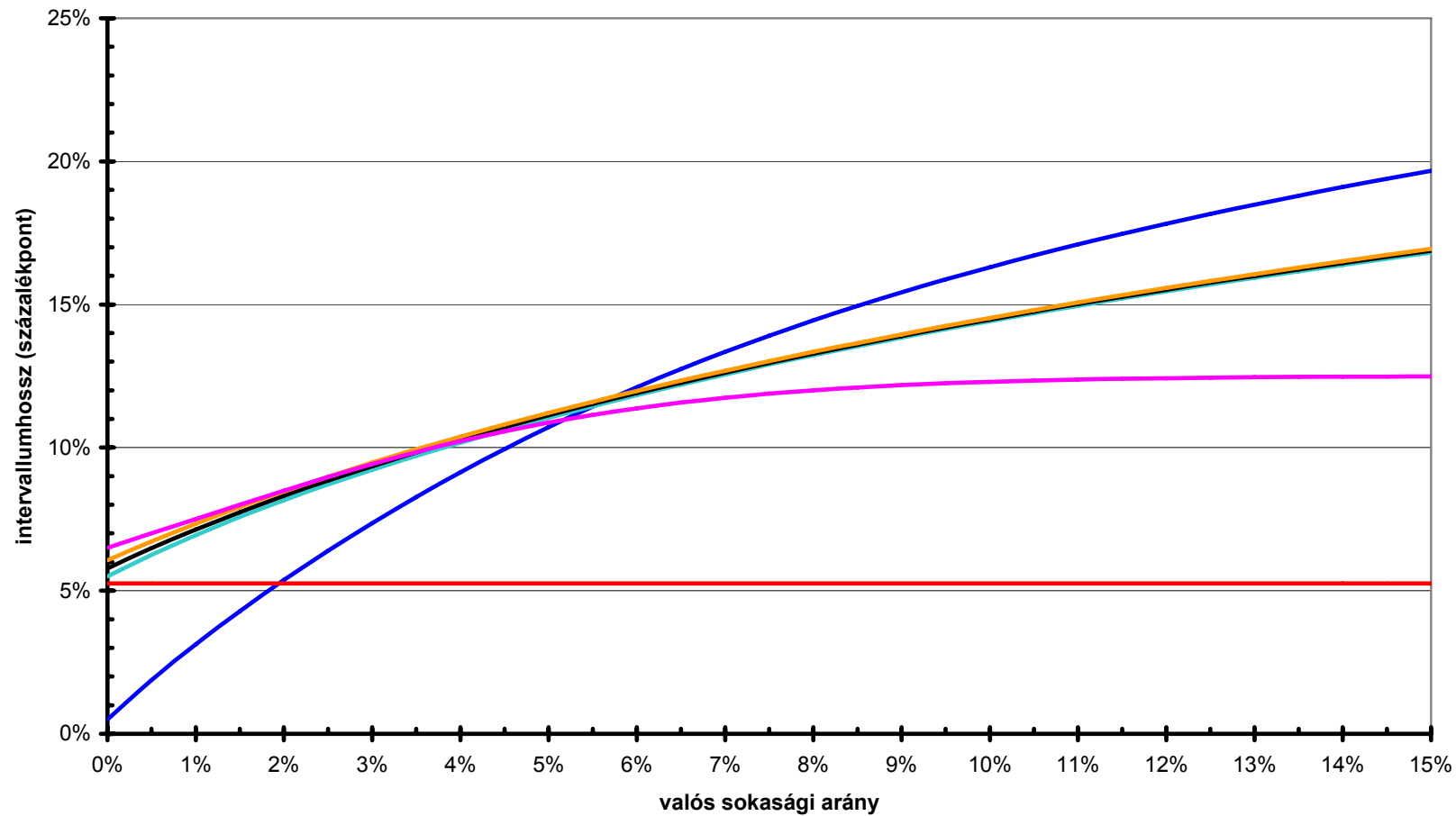


N =	200
n =	50
p =	5,00%

M1
M2
M3
B1
B2
MxB1

20. ábra: Az elemzés I. outputja a (200/50/5%) forgatókönyv esetén

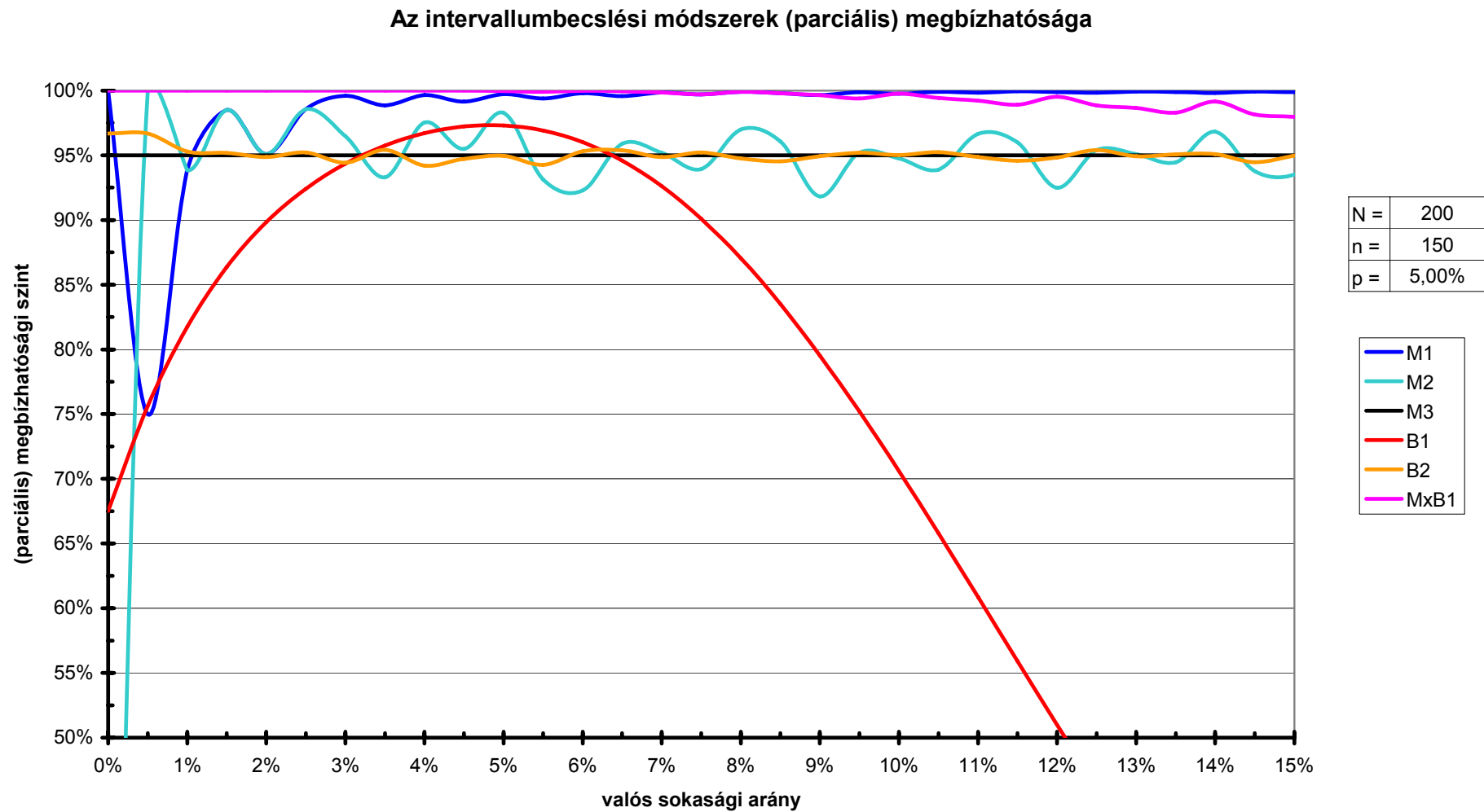
Az intervallumbecslési módszerek várható intervallumhossza



N =	200
n =	50
p =	5,00%

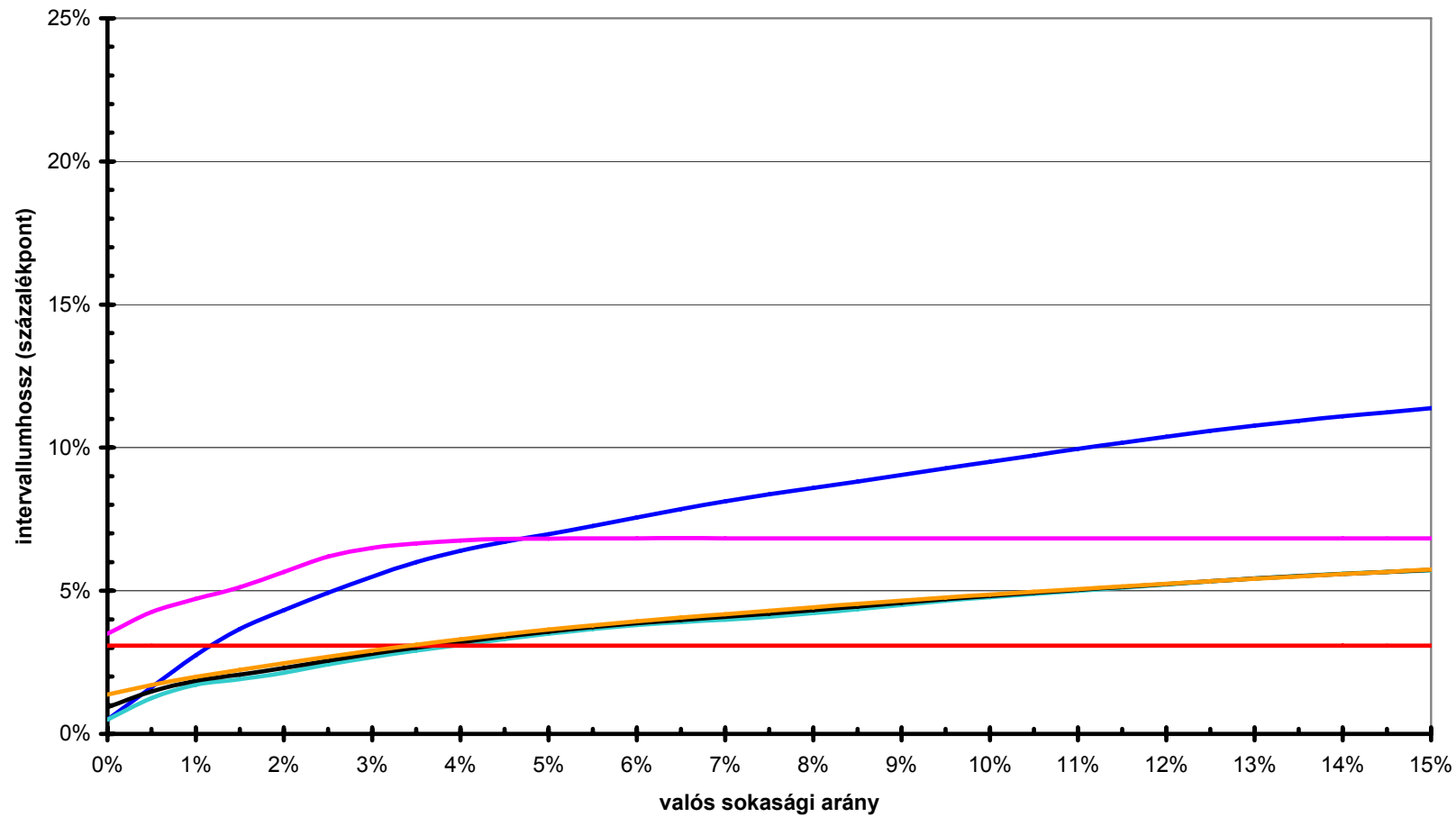


21. ábra: Az elemzés II. outputja a (200/50/5%) forgatókönyv esetén



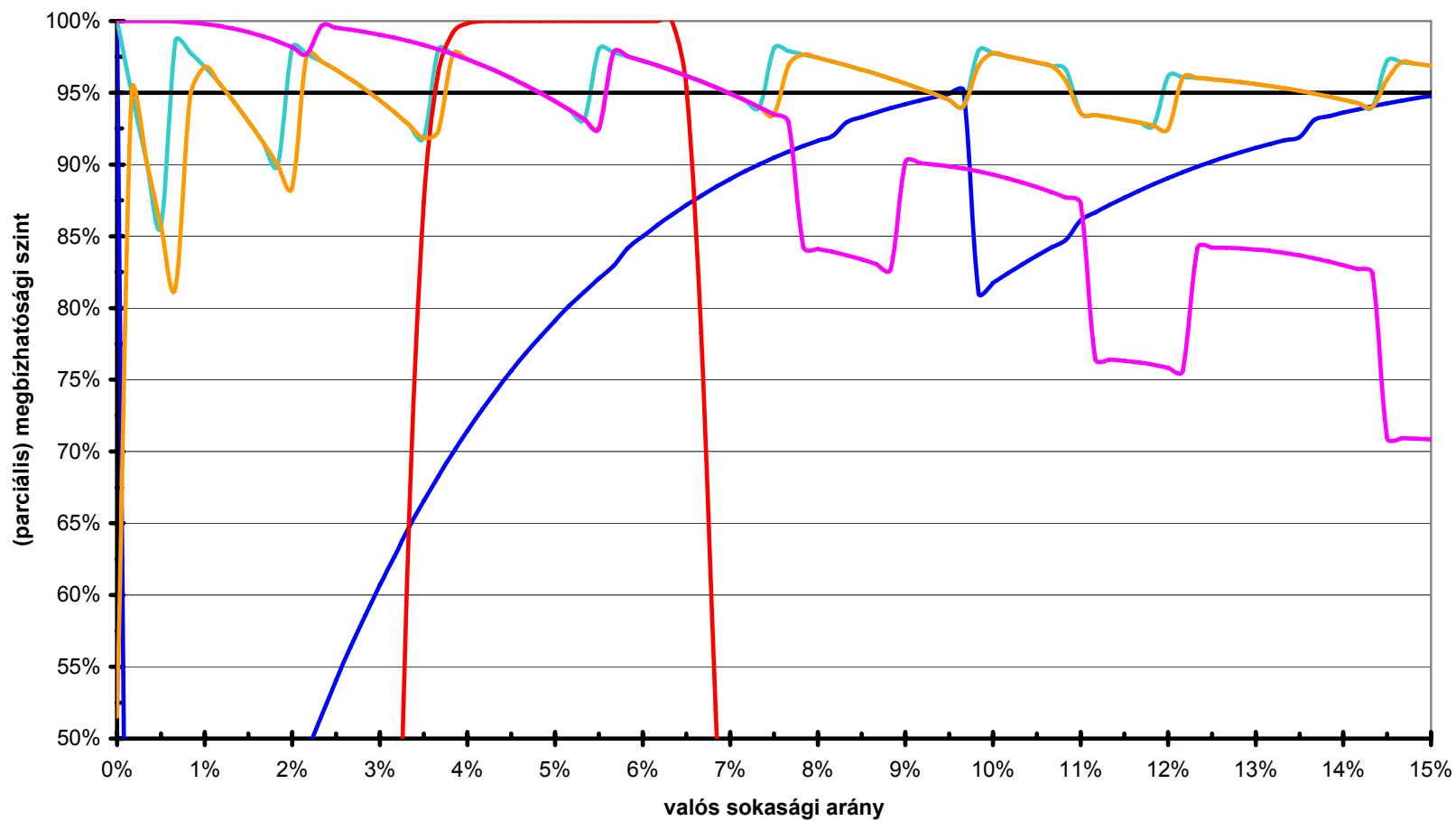
22. ábra: Az elemzés I. outputja a (200/150/5%) forgatókönyv esetén

Az intervallumbecslési módszerek várható intervallumhossza



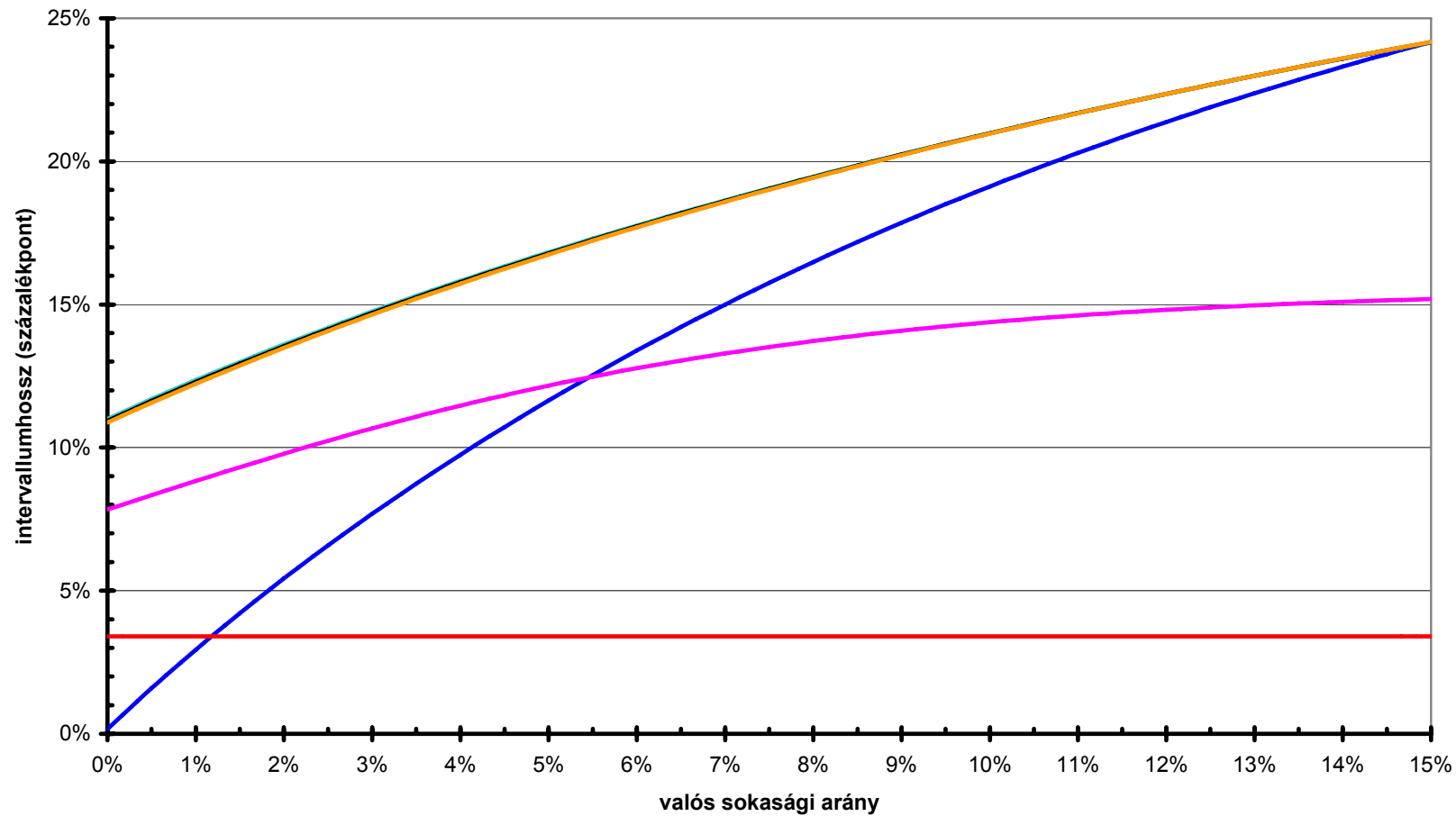
23. ábra: Az elemzés II. outputja a (200/150/5%) forgatókönyv esetén

Az intervallumbecslési módszerek (parciális) megbízhatósága



24. ábra: Az elemzés I. outputja a (600/30/5%) forgatókönyv esetén

Az intervallumbecslési módszerek várható intervallumhossza

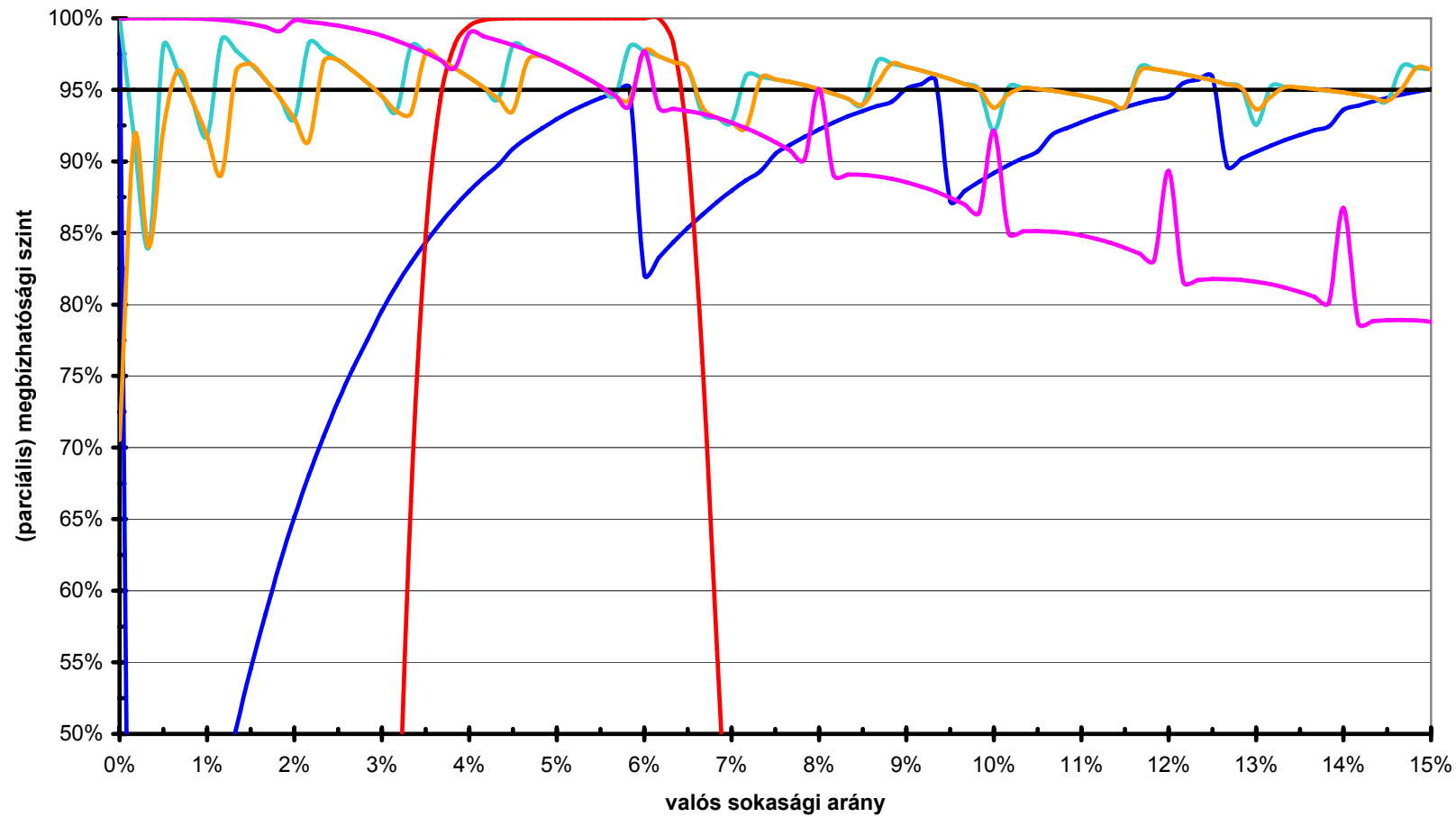


N =	600
n =	30
p =	5,00%



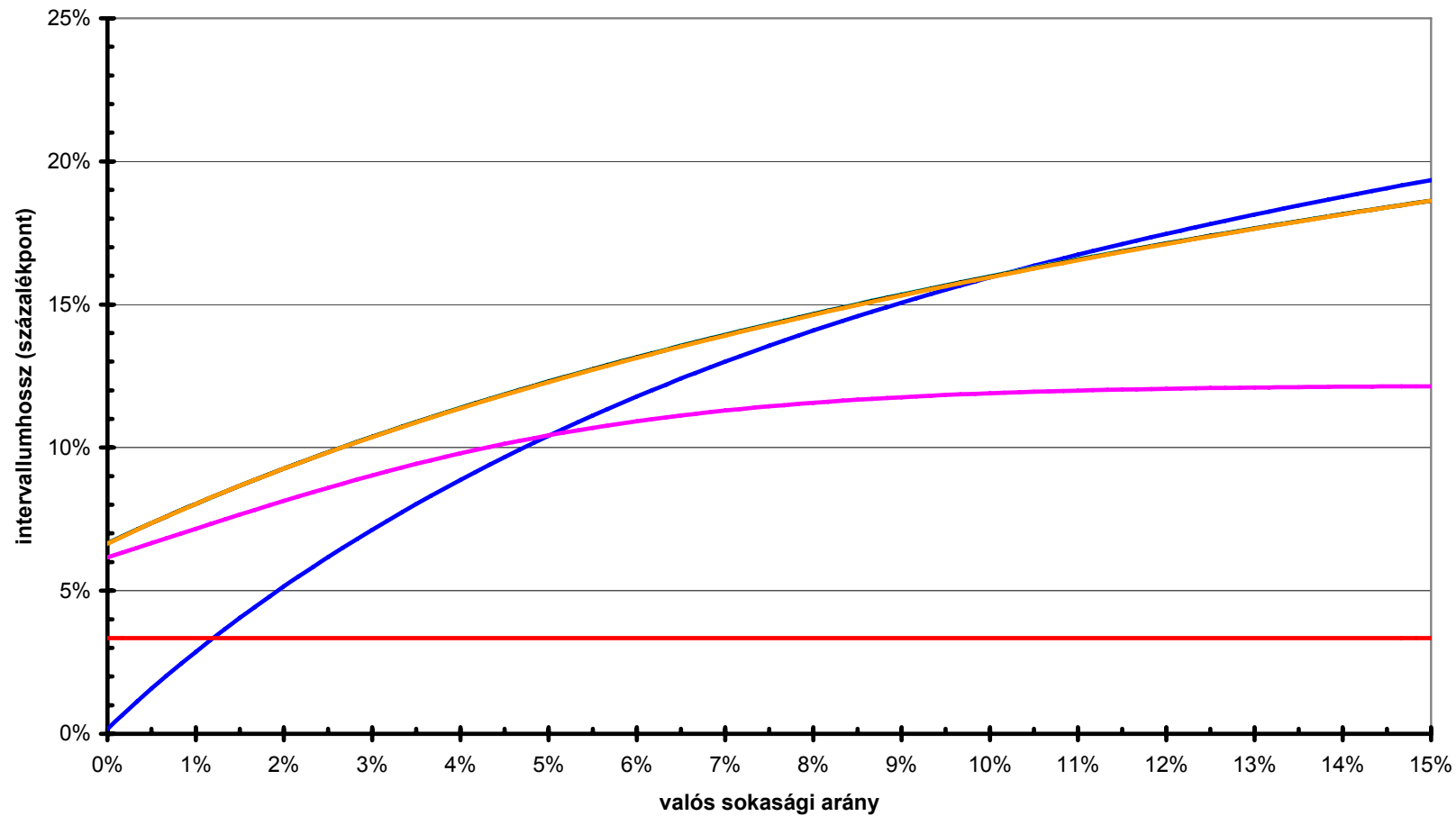
25. ábra: Az elemzés II. outputja a (600/30/5%) forgatókönyv esetén

Az intervallumbecslési módszerek (parciális) megbízhatósága



26. ábra: Az elemzés I. outputja a (600/50/5%) forgatókönyv esetén

Az intervallumbecslési módszerek várható intervallumhossza

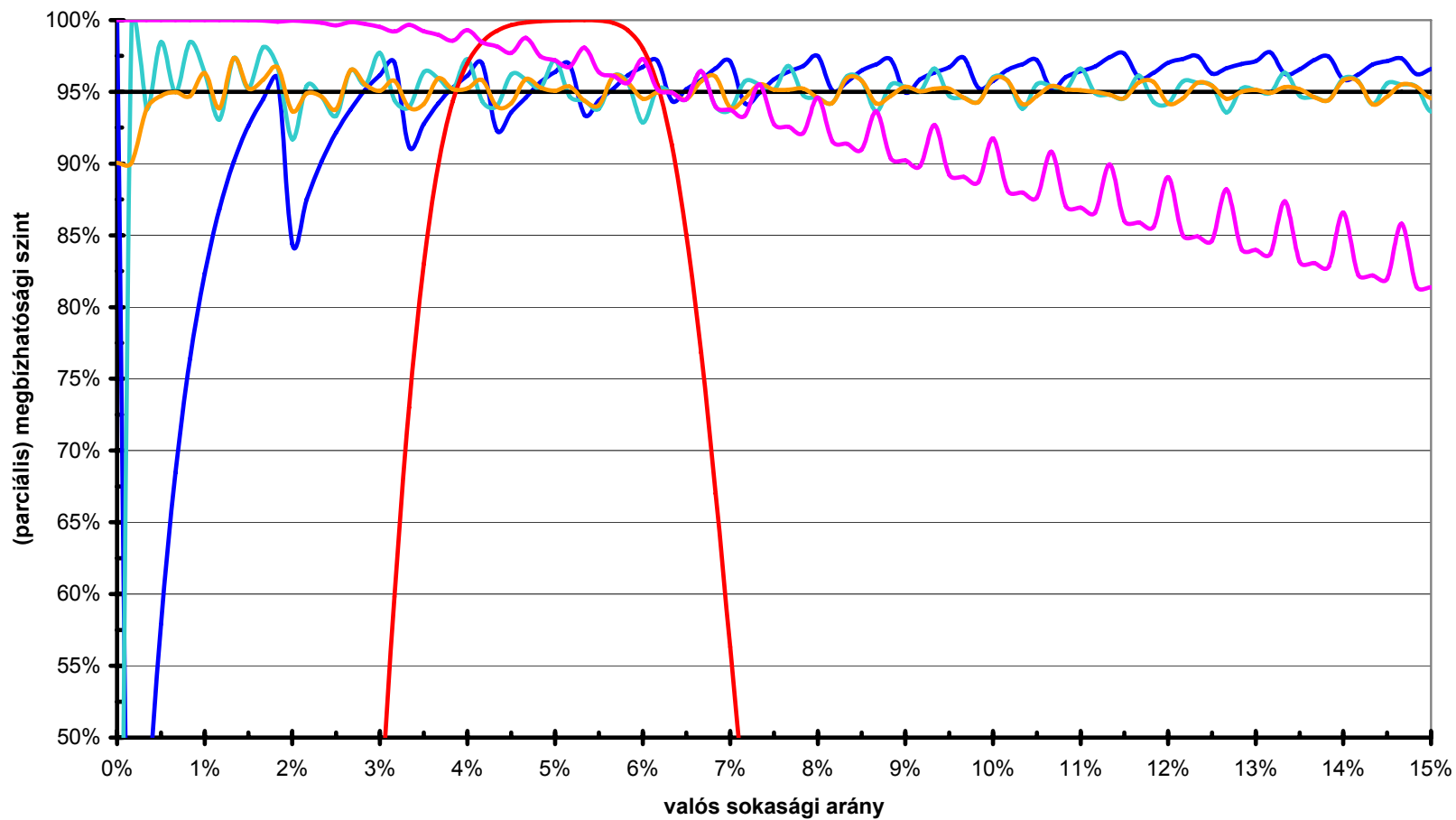


N =	600
n =	50
p =	5,00%

M1
M2
M3
B1
B2
MxB1

27. ábra: Az elemzés II. outputja a (600/50/5%) forgatókönyv esetén

Az intervallumbecslési módszerek (parciális) megbízhatósága

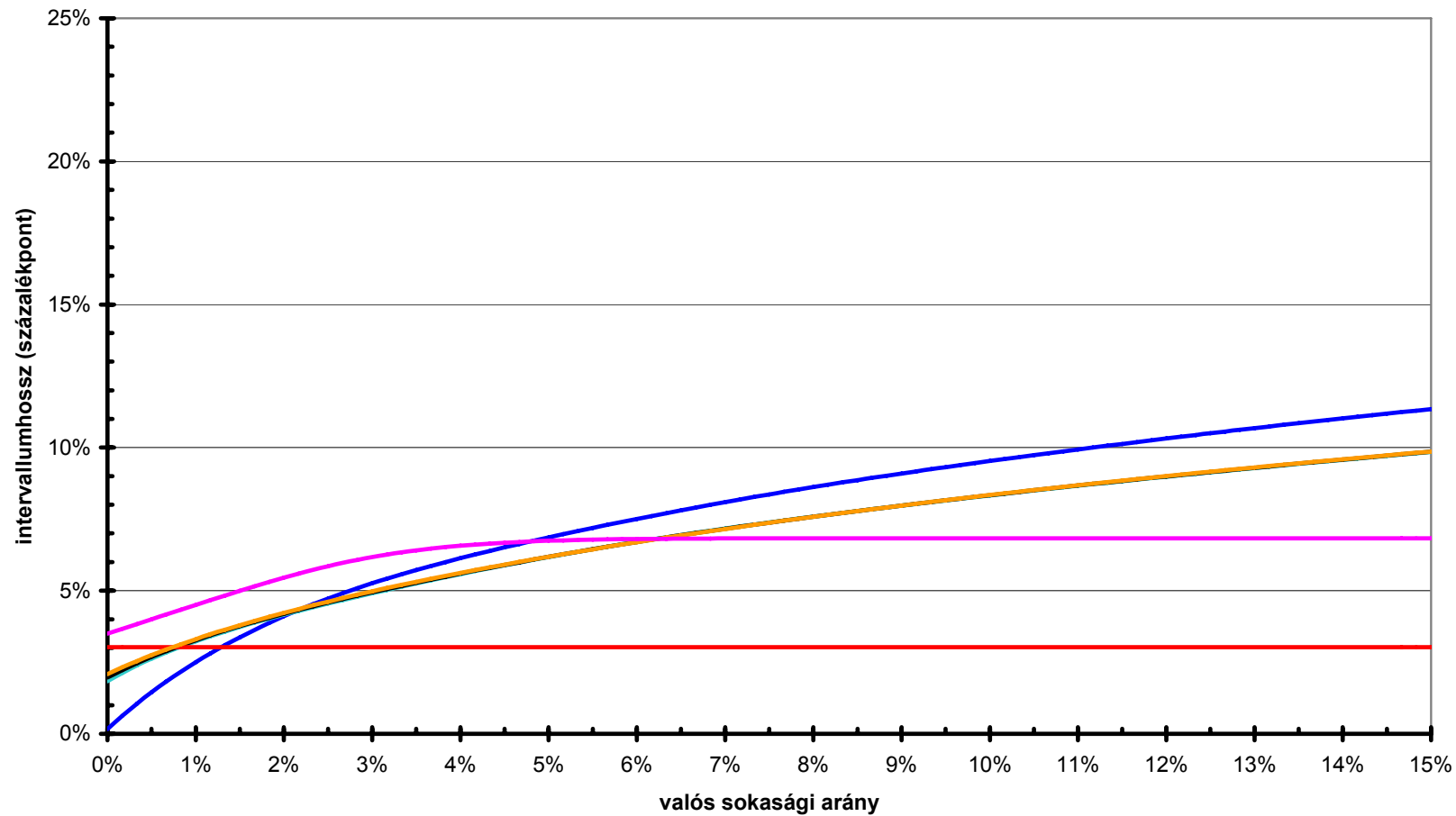


N =	600
n =	150
p =	5,00%

M1
M2
M3
B1
B2
MxB1

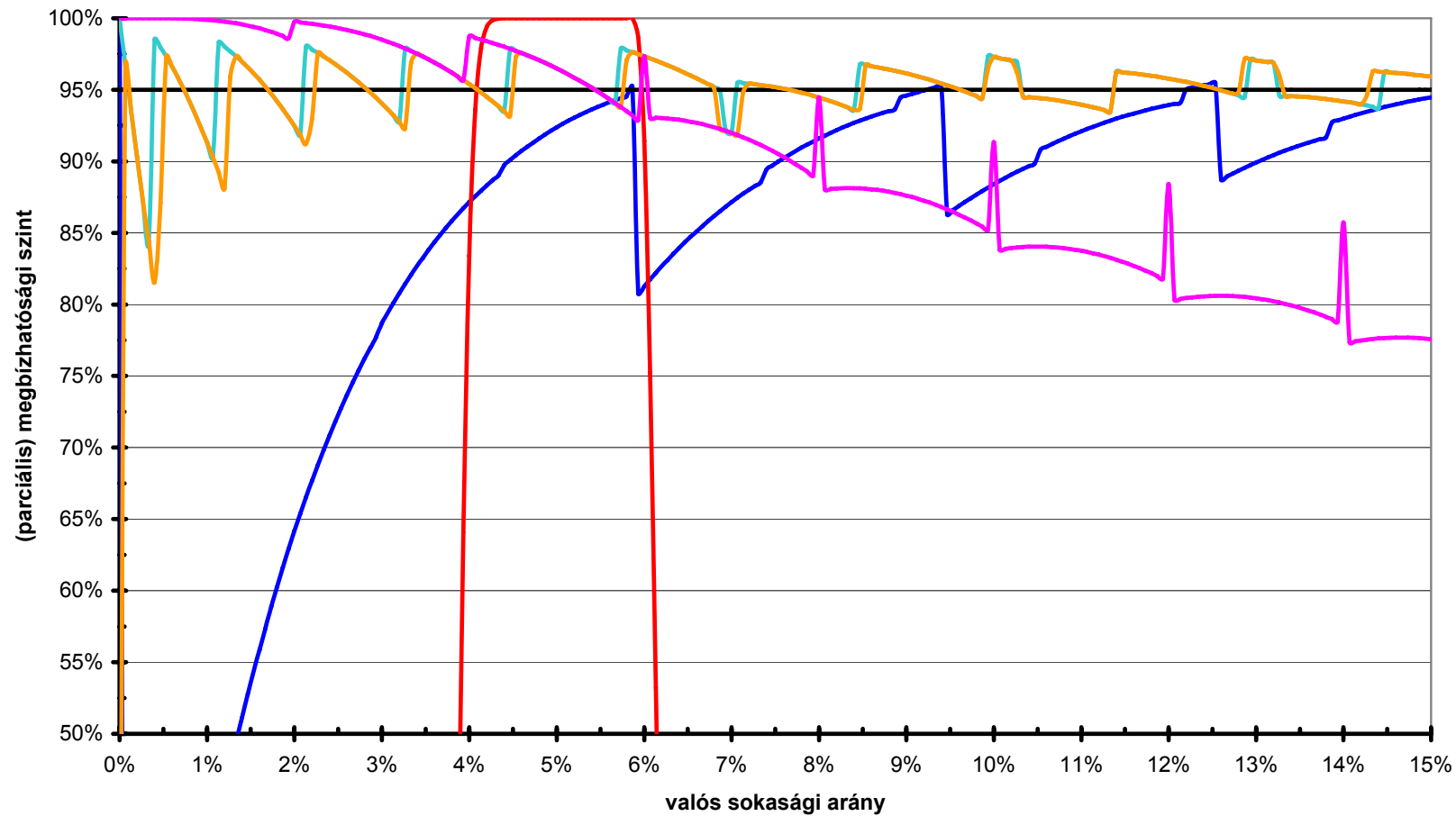
28. ábra: Az elemzés I. outputja a (600/150/5%) forgatókönyv esetén

Az intervallumbecslési módszerek várható intervallumhossza



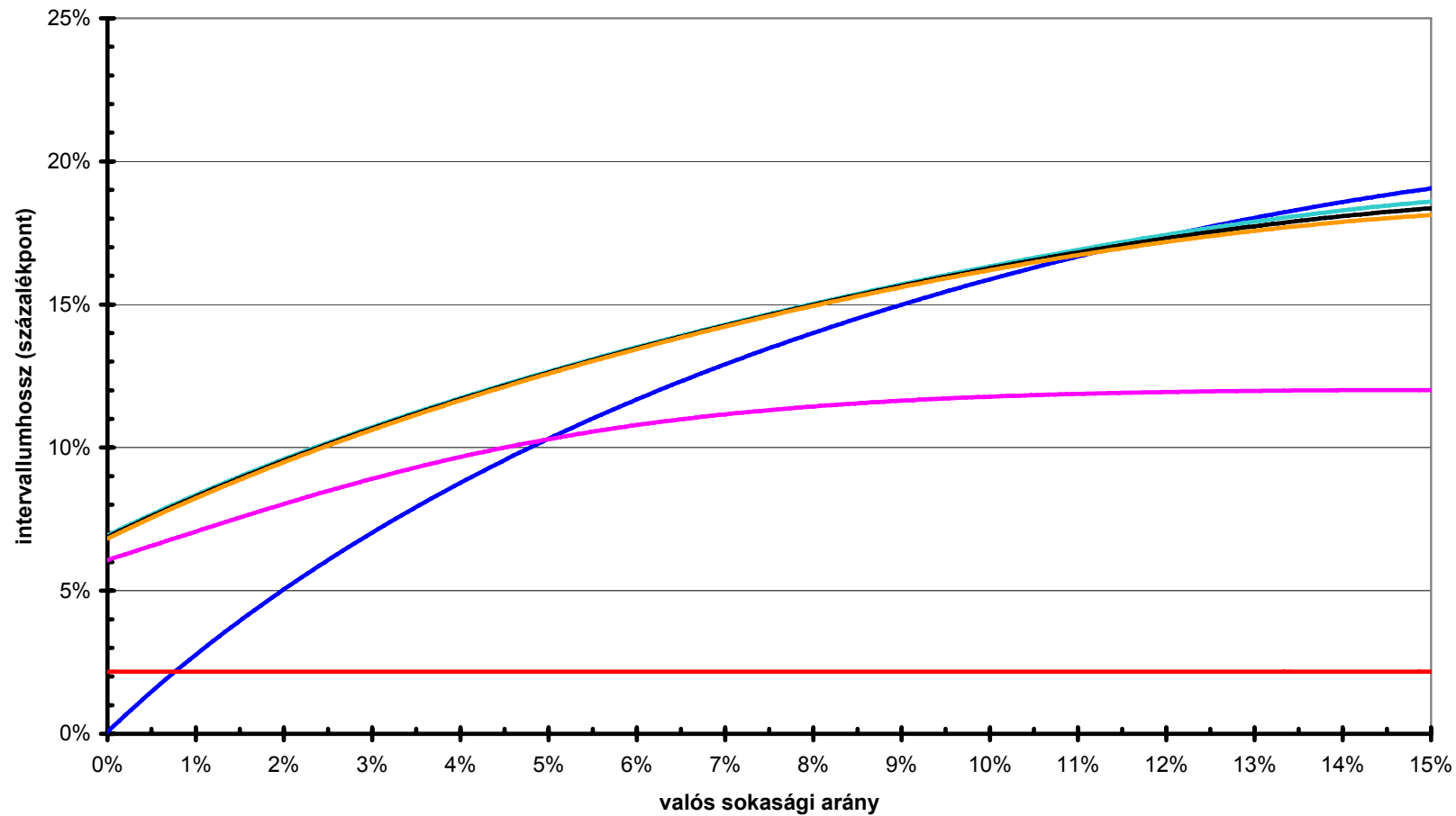
29. ábra: Az elemzés II. outputja a (600/150/5%) forgatókönyv esetén

Az intervallumbecslési módszerek (parciális) megbízhatósága



30. ábra: Az elemzés I. outputja az (1500/50/5%) forgatókönyv esetén

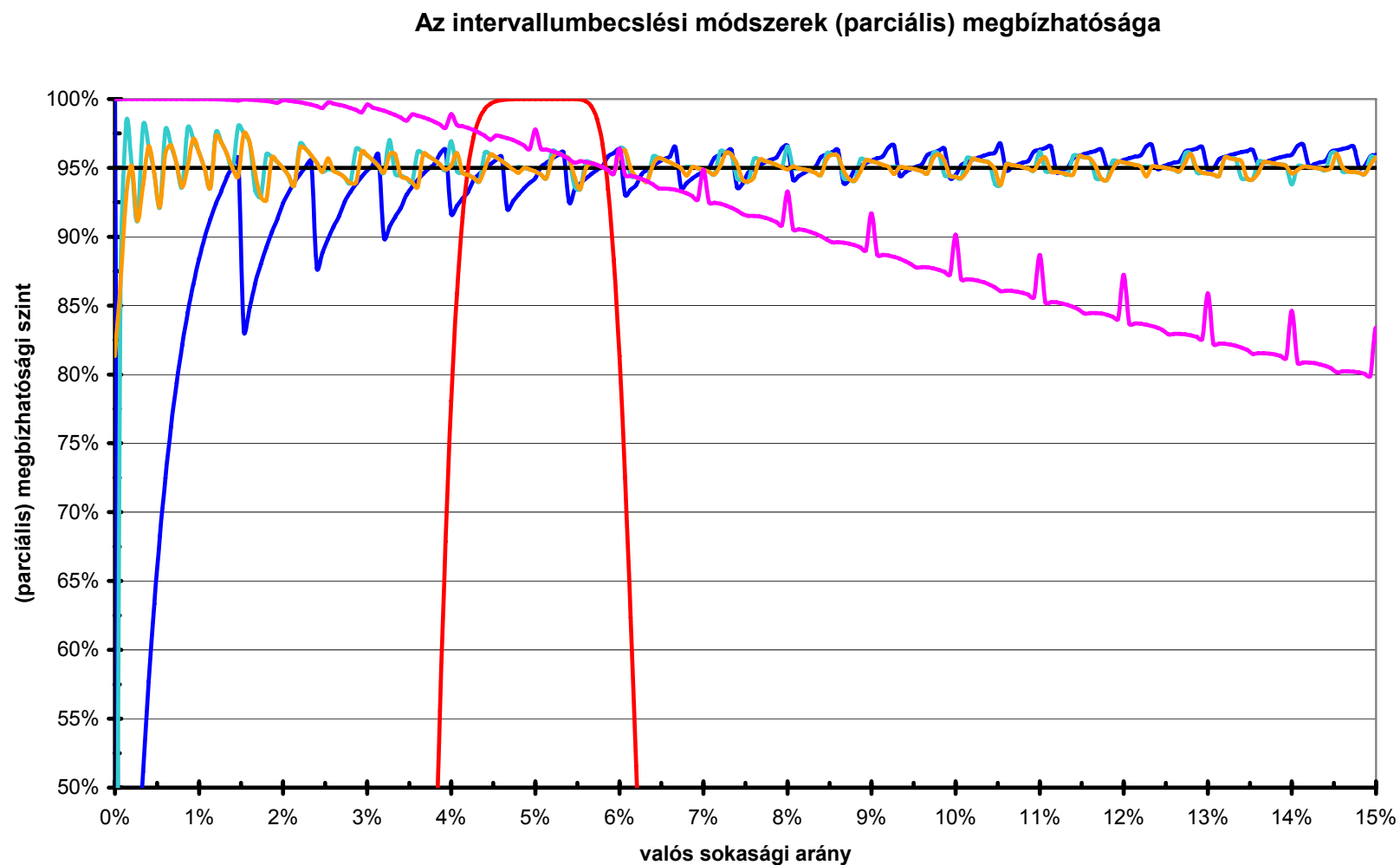
Az intervallumbecslési módszerek várható intervallumhossza



N =	1500
n =	50
p =	5,00%

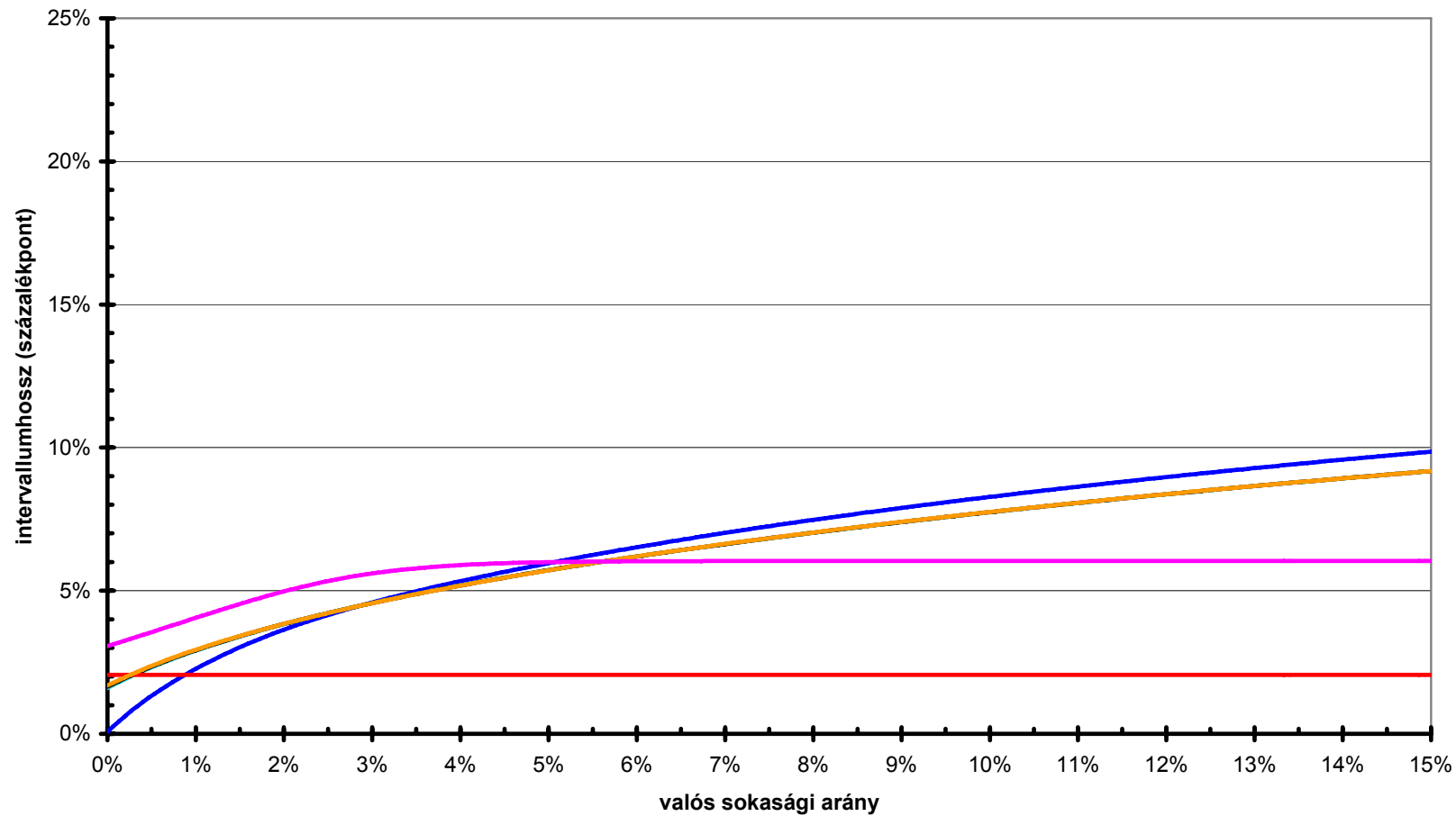
M1
M2
M3
B1
B2
MxB1

31. ábra: Az elemzés II. outputja az (1500/50/5%) forgatókönyv esetén



32. ábra: Az elemzés I. outputja az (1500/200/5%) forgatókönyv esetén

Az intervallumbecslési módszerek várható intervallumhossza



N =	1500
n =	200
p =	5,00%

M1
M2
M3
B1
B2
MxB1

33. ábra: Az elemzés II. outputja az (1500/200/5%) forgatókönyv esetén