

EÖTVÖS LORÁND TUDOMÁNYEGYETEM

BUDAPESTI CORVINUS EGYETEM

Gajdos Zsuzsanna

VAR BACKTESTING KLASSZIKUS ÉS BAYESI STATISZTIKAI MÓDSZEREKKEL

DIPLOMAMUNKA

Biztosítási és pénzügyi matematika MSc.

Kvantitatív pénzügyek specializáció

Témavezető: Szádoczki Zsombor



ELTE
EÖTVÖS LORÁND
TUDOMÁNYEGYETEM



Budapest, 2024

Köszönetnyilvánítás

Először is szeretném megköszönni a témavezetőmnek, Szádoczki Zsombornak a rengeteg segítséget és az alapos munkát, valamint a kaotikus napirendemmel szembeni megértést és türelmet is.

Szeretném továbbá megköszönni a barátomnak és a családomnak a feltétel nélküli támogatást és szeretetet. Köszönettel tartozom szaktársamnak, Áginak is, hogy mindig ott volt nekem a nehéz időszakokban és nem hagyta, hogy feladjam.

Hálás vagyok továbbá a főnökömnek, Miklósnak, amiért megismertetett a backtesting világgal és válaszolt a kismillió kérdésemre a megértéshez vezető úton.

Tartalomjegyzék

Köszönetnyilvánítás	1
Ábrák jegyzéke	3
Táblázatok jegyzéke	4
1. Bevezetés	5
2. Irodalomáttekintés	7
3. VaR backtesting	9
3.1. Piaci kockázatok	10
3.2. Partnerkockázat	15
4. Binomiális teszt	18
5. Bayesi statisztikán alapuló backtesting	23
5.1. Bevezetés, motiváció	23
5.2. Példa a volatilitás félrespecifikálásra	24
5.3. Prior szerepe	26
5.4. A prior és posterior cseréje	29
5.5. Hipotézisvizsgálat	32
6. Tesztek	34
6.1. Első scenárió	34
6.2. Második scenárió	36
6.3. Eredmények összegzése	39
7. Konklúzió, további kutatási kérdések	40
Hivatkozásjegyzék	42

Ábrák jegyzéke

1.	A Binom(250; 0,01) eloszlás értékei néhány adatpont esetén	13
2.	A Binom(250; 0,01) eloszlásfüggvénye és az értékhez tartozó zóna	13
3.	A különböző időtávok melletti sampling (saját szerkesztés, forrás: [7]) . . .	17
4.	A különböző percentilisekhez tartozó kritikus értékek mintaelemszám függvényében	18
5.	Az eljárásból fakadó elsőfajú hiba mintaelemszám függvényében	19
6.	Az eljárásból fakadó és a szimulált elsőfajú hiba mintaelemszám függvényében	20
7.	Az elsőfajú hiba nagysága a szimulált adatmennyiség függvényében, több percentilis együttes használata mellett	21
8.	Példa a prior és posterior eloszlásokra	24
9.	Példa a prior és posterior eloszlásokra	25
10.	A posterior eloszlás várható értékének és a kumulált tapasztalati szórásnak a hányadosa adatmennyiség függvényében	26
11.	1. Példa posterior függvényei Gamma(10; 10) és Uniform(0,5; 2) eloszlású priorral	27
12.	1. Példa posterior függvényei Uniform(0,5; 1,5) eloszlású priorral	27
13.	Bayesi backtesting, ha a keresett paraméter nincsen a prior értékkészletében	28
14.	A kumulált realizált adatok szórása és a különböző mintanagyságokon update-elt posteriorok várható értéke	31
15.	1. Példa a posterior percentiliseivel	33
16.	Próba erő összehasonlítása az első scenárió mellett	35
17.	A második scenárió realizációinak hisztogramja	37
18.	Próba erő összehasonlítása a második scenárió mellett	38

Táblázatok jegyzéke

1. A piaci kockázatok backtesting feltevései melletti valószínűségek és elsőfajú hibák, lásd: ([4], Table 1.) 12
2. A piaci kockázatok backtesting eredményei alapján meghatározott büntetőfaktora 15

1. Bevezetés

A pénzügyi intézmények biztonságos működésének nemzetközi szabályozása az 1970-es évektől a Bázeli Bankfelügyeleti Bizottság (Basel Committee on Banking Supervision) ajánlásaira támaszkodik. A bankok kockázatelemzési módszereinek folyamatos fejlődése azt eredményezte, hogy 2004-ben megjelent a *Revised Capital Framework* ([3]), ami a Bazel II néven terjedt el világszerte.

A Bazel II három fő pillérre bontotta a kockázatalapú tőketartalékolás keretrendszerét:

- minimum tőkekövetelmény meghatározása különböző kockázati típusok esetén,
- a felügyelet intézményesítése,
- nyilvánosságra hozatali követelmény.

A pénzügyi intézmények kockázatainak kiterjedt besorolását a 2010-ben közzétett *A global regulatory framework for more resilient banks and banking systems*, azaz a Bazel III ([6]) tartalmazza átfogóan. Ebben a Bazel II által is kezelt piaci, hitel- és működési kockázatok mellett már a likviditási- és rendszerkockázat is megjelent.

A kockázatok utáni tőkekövetelmény számolására két lehetőség közül választhatnak a pénzügyi intézmények: a sztenderd módszert vagy a belső modelleken alapuló módszert. Míg a sztenderd módszer a szabályozók által megadott, szigorú lépéssorozaton alapszik, addig a belső modell alapú megközelítés esetén a hitelintézetek a saját maguk által meghatározott és kalibrált veszteségeloszlás alapján számolhatják a szavatolótőke-szükségletüket.

A veszteségeloszlást a különböző kockázati faktorok segítségével szimulálják, majd ezen eloszlás megfelelő percentilise, az ún. kockázatosított érték, a VaR (Value-at-Risk) megfelelő átskálázásával kapják a tőkeszükségletüket.

A felügyelet részéről jogos tehát a kérdés, miszerint ha a saját modellt használják a pénzintézetek a tőkekövetelményük meghatározásához, akkor biztosan megfelelően vannak-e kalibrálva és felépítve a modelljeik. Másfelől az egyes hitelintézetek számára is fontos, hogy az adott tőkemennyiség elegendő-e az esetleges veszteségek fedezésére.

Ezen cél mentén adta ki a Bázeli Bankfelügyelet a *Sound practices for backtesting counterparty credit risk models* ([7]) című publikációt, ami pontos keretrendszert biztosított a prediktív modellek minőségi ellenőrzésére.

A backtesting matematikai szempontból egy olyan retrospektív eljárás, amivel egy adott időhorizonthoz tartozó szimulált eloszlás idősorát a realizált értékekkel összevetve megkapjuk, hogy az adott szimuláció helyes volt-e adott szignifikanciaszint mellett.

Statisztikai szempontból a backtesting az adatok hiánya miatt jelent kihívást: például egy olyan kockázati faktor esetén, amely szimulálása csak fél éve történik, az egyhetes horizont melletti időszora csak 26 adatpontot tartalmaz. A bázeli keretrendszer az adatmennyiség alsó határát 20 adatpontnál határozta meg. Emellett a módszertan diszkrét volta is azt eredményezi, hogy elsőfajú és másodfajú hiba szempontjából a kapott eredmények szignifikanciáját nem lehet biztosítani.

A hipotézisvizsgálatok során megadott konfidenciaintervallum egyértelműen meghatározza az elsőfajú hibát - elméletben. A kevés adaton végzett vizsgálat során azonban a kalibrált és a ténylegesen megjelenő elsőfajú hiba nagy mértékben eltérhet. Ez azt eredményezheti, hogy megnő azon esetek száma, amikor a jól működő modelleket elutasítjuk, ami feleslegesen megnöveli a szavatolótőke mennyiségét.

A próba ereje azért fontos, mert magas másodfajú hiba esetén nem utasítjuk el a rosszul kalibrált modelleket, ami alultőkésítéshez vezethet, ezáltal megjelenik egy olyan kockázat, amit nem fedezünk.

Erre a problémára jelenthet egy megoldást a bayesi statisztikán alapuló backtesting, ami erő és elsőfajú hiba szempontjából gyorsabban konvergál az ideális hipotézisvizsgálathoz adatmennyiség függvényében, mint a klasszikus, [15]-ben is említett eljárás.

Ezen dolgozat az egyik jelenleg is elterjedt, autokorreláció nélküli backtesting eljárást hasonlítja össze a Zelvyte és Arnsdorf által [16]-ban bemutatott módszertannal elsőfajú hiba és próba erő szempontjából. Továbbá néhány példán keresztül illusztrálja, hogy a különböző módon félrespecifikált modelleket miként lehet hatékonyan felismerni ezen módszertant felhasználva.

2. Irodalomáttekintés

A jelenlegi szabályozás alapját az 1974-ben alapított Bázeli Bankfelügyeleti Bizottság (Basel Committee on Banking Supervision, BCBS) ajánlásai adják.

Az első kidolgozott ajánlást 1988-ban adták ki *International Convergence of Capital Measurement and Capital Standards* ([1]) néven, ami Bazel I-ként vált ismertté. Ezen ajánlás csak a hitelkockázatot vette figyelembe, valamint csak kezdetlegesen mérte fel a különböző kockázatok súlyát a veszteségeloszlás meghatározásánál.

Jól mutatja a pénzüintézetek igényét a kockázatuk pontos ismeretére az, hogy a szabályozás következő lépcsőfoka előtt, 1992-ben, négy év fejlesztés után a JP Morgan szakemberei kiadták *RiskMetrics - Technical Document* ([9]) néven a modern kockázatelemzés alapjait adó módszertant. Kidolgozták többek között a kockázatot érték (Value-at-Risk, VaR) alapú kockázatmérést a hitelkockázat mellett a piaci kockázatra is, valamint kitértek a modellek ellenőrzésére szolgáló módszerre, a backtestingre is.

A szabályozók a piaci helyzetre reagálva 1996-ban beépítették a Bazel I módszertanába a piaci kockázatokat és a VaR alapú megközelítést, ez lett az *Amendment to the Capital Accord to Incorporate Market Risks* ([2]).

A szabályozók két lehetséges megoldást kínálnak a pénzüintézeteknek: a sztenderd módszert és a belső módszereken alapuló tőkeszükséglet-számítást. Míg a sztenderd módszer konkrét lépéseken keresztül határozza meg a szükséges szavatoló-tőke mennyiségét, addig a belső modellen alapuló megközelítés esetén a pénzüintézeteknek lehetősége van modellezni, becsülni és kalibrálni a saját veszteségeloszlásukat a [2, Part B.4] által meghatározott minimumkövetelmények mellett.

A szabályozás következő lépcsőfoka a 2004-ben kiadott Bazel II, azaz a *Revised Capital Framework* ([3]), amit 2009-ben a 2008-as gazdasági válság miatt kiegészítettek (ez lett az ún. Bazel 2.5 ([5])) többek között a stresszelt VaR-ral és annak backtesting követelményével.

A Bazel II-es ajánlás kiegészítéseként jelent meg a *Supervisory framework for the use of 'backtesting' in conjunction with the internal models approach to market risk capital requirements* ([4]), amiben a backtesting ma is használt red-amber-green / traffic light megközelítése is szerepel. Ezen módszertan a backtestingről szóló kiegészítés ([2, Part B.4] (j) pontjának konkretizálása.

A következő állomás a 2010-ben kiadott Bazel III, azaz a *A global regulatory framework for more resilient banks and banking systems* ([6]) volt, ami már a partnerkockázatot (counterparty credit risk), valamint az azzal kapcsolatos backtestinget is újraértékelte, ez lett a *Sound practices for backtesting counterparty credit risk models* ([7]). A Bazel III 117.46i

pontja szerint nem csak a piaci kockázati faktorokat, hanem hipotetikus trade-eket és a pénzintézet reprezentatív portfóliói által meghatározott jövőbeli várható veszteségeloszlást is backtestelni kell.

A szabályozó irodalom és a RiskMetrics mellett érdemes még megemlíteni a szabályozás javítására, pontosítására megjelent új módszertanokat is. Ilyen Christoffersen 1998-as cikke ([12]), ami az intervallumbecslés minőségi vizsgálatát tárgyalja matematikai megközelítésből. Érdemes megemlíteni Berkowitz 2001-es cikkét ([13]) is, ami Christoffersenhez hasonlóan matematikai megközelítéssel áll a témához, azonban a gyakorlatban megjelenő problémára, az alacsony adatszámra reflektálva dolgoz ki egy eljárást.

A Bazel III-as szabályozás újításait a backtestinggel kapcsolatban Anfuso, Karyampas és Nawroth 2013-as tanulmánya ([11]) bontotta ki részletesen.

Ezen dolgozatban tárgyalt bayesi statisztikán alapuló módszertant Zelvyte és Arnsdorf dolgozták ki ([16]), és ezzel kínáltak egy megoldást a klasszikus hipotézisvizsgálat p -értékével kapcsolatos problémára. Ezen módszertan nem csak a hipotézisvizsgálat minőségét javítja (elsőfajú hiba és próba erő), hanem a realizált adatok és a prior eloszlás alapján a modellek pontos kalibrálásában is segít.

3. VaR backtesting

Ahhoz, hogy a pénzüzetek belső modellek használatával határozzák meg a tőkekövetelményüket, számos, minőségi követelménynek kell megfelelniük. Ilyen többek között a piaci kockázatokból származó veszteségeloszlás monitorozása is, amit

- naponta,
- 99%-os, egyoldali konfidenciaintervallummal,
- minimum 10 napos tartási periódust feltételezve,
- legalább egy éves megfigyelési időszak mellett szükséges modellezni.

A szabályozás nem írja elő, hogy milyen modelleket kell alkalmazni, azonban a Risk-Metrics ([9]) óta a kockázatotott érték (VaR) alapú megközelítés az elterjedt. A VaR a veszteségeloszlás egy percentilise, amit adott időtáv és konfidenciaintervallum függvényében tudunk pontosan meghatározni. A kockázatotott érték, mint kockázati mérték, tehát azt mutatja meg, hogy mi az a mennyiség, aminél nem fogunk nagyobb veszteséget realizálni egy adott időtávon, egy adott valószínűség mellett.

Érdeemes kitérni a veszteség, mint pénzügyi fogalom lehetséges értékeire annak függvényében, hogy a kockázatkezelés mely területéről tekintünk rá. A két legnagyobb kockázattípus, amit a szabályozás is megkülönböztetetten kezel, az a piaci és a hitelkockázat.

A piaci kockázatok területén a kockázat abból fakad, hogy a piaci árak bármely irányba elmozdulhatnak, ezáltal a portfóliónk értéke a jövőben lehet pozitív és negatív is. A kockázat így azt a bizonytalanságot jelenti, miszerint a jövőben lehet nyereségünk és veszteségünk is. Innen ered az elnevezés: P&L függvény.

A hitelkockázat azonban abból a jövőbeli bizonytalanságból fakad, hogy nem tudjuk, a partnerünk fizetőképese lesz-e vagy sem. Tehát a kockázatkezelés ezen területén a legjobb kimenetel az, ha senki nem válik fizetéseképtelenné, így a hitelkockázatok veszteségfüggvénye mindig nemnegatív.

Formálisan, ha az X_τ véletlen változó a veszteségeink lehetséges értékei a τ időpontban, akkor a

$$\text{VaR}_\alpha(X_\tau) = \inf\{x : P(X_\tau \leq x) \geq \alpha\}, \quad (1)$$

ahol $\alpha \in [0; 1]$, τ pedig rögzített.

A VaR azért is tudott elterjedni, mert egy mennyiségre egyszerűsítette a veszteségeloszlást. Azonban amellet, hogy nem egy koherens kockázati mérték ([18]), a percentilisen túli

veszteségeket sem veszi figyelembe. Erre kínált megoldást a 2012-ben bevezetett FRTB (*Fundamental review of the trading book* ([8])), ami a VaR-on túli veszteségeket is figyelembe véve egy új kockázati mérték, az Expected Shortfall használatát javasolja. Az FRTB implementálása a gyakorlatba azonban olyan komplexitásbeli ugrás volt a pénzüzetek számára, hogy a szabályozás a mai napig nem várja el ezen framework használatát¹.

A backtesting tehát annak az ellenőrzése a múltbeli adatok alapján, hogy az előre meghatározott, szimulált percentilis (VaR) megegyezik-e a valós eloszlás percentilisével, azaz a kockázati mérték, ami alapján a pénzüzetek tőkét allokálnak, helyes-e. Ahhoz, hogy pontos képünk legyen a valós eloszlásunkról, több percentilis egyidejű figyelembevételére lenne szükség, azonban kockázatelemzési szempontból csak a felső percentilisek számítanak, hiszen azok alulbecslése jelent szavatolótőkével nem fedezett kockázatot.

A Bázeli II backtestingre vonatkozó kiegészítésének ([4]) II. fejezete elismeri, hogy a szélső percentilisek becslése nehezen ellenőrizhető megbízhatóan, mégis a szabályozás része, hiszen a Bázeli II-es elsődleges szavatolótőke-számítás alapját adja. Innen adódik tehát a backtesting, mint statisztikai probléma: hogyan lehet a múltbeli adatok alapján eldönteni, hogy a szimulált eloszlásunk megfelelt-e a valóságnak.

3.1. Piaci kockázatok

A Bázeli Bankfelügyeleti Bizottság által kínált módszer egyszerű heurisztikán alapul: a VaR-on túli realizációk számát elosztva az összes realizációval a VaR percentilisét kell kapnunk. Például 200 kereskedési napból a $VaR_{0,99}$ összesen $200 \cdot (1 - 0,99) = 2$ realizációt nem fedez a percentilis definíciója alapján.

A Bázeli II-es szabályozás ([4]) III.(a) pontja három zónát határoz meg a következők alapján:

- legyen N a megfigyeléseink száma,
- $(R_i)_{i=1}^N$ a veszteségek realizációi, valamint
- $(VaR_\alpha(i))_{i=1}^N$ a becsült veszteségeloszlások α percentiliseinek sorozata.

Legyen a VaR által nem fedezett (azaz a kockázatosított értéknél nagyobb) veszteségek indikátorváltozója a következő:

$$I_t = \begin{cases} 1, & \text{ha } R_t \geq VaR_\alpha(t), \\ 0, & \text{különben.} \end{cases}$$

¹<https://www.numerix.com/blog/frtb-are-you-ready-for-a-new-era-in-market-risk-management>

A backtesting során azt akarjuk ellenőrizni, hogy a VaR-hoz tartozó α megegyezik-e a

$$\hat{\alpha} = \frac{1}{N} \sum_{t=1}^N I_t$$

statisztikával. A Bizottság által meghatározott három zóna a következők szerint alakul:

$$\sum_{k=0}^{\sum_{t=1}^N I_t} \binom{N}{k} p^k (1-p)^{N-k} = \alpha = \begin{cases} \text{red zone,} & \text{ha } \alpha \geq 0,9999, \\ \text{yellow zone,} & \text{ha } \alpha \geq 0,95, \\ \text{green zone} & \text{ha } \alpha < 0,9. \end{cases}$$

Vegyük észre, hogy az egyenlet bal oldala az $(N; p)$ paraméterű binomiális eloszlás eloszlásfüggvényének értéke a nem fedezett kockázatok számánál.

Ezen megközelítés feltétele tehát, hogy a becsült eloszlásaink és a realizációink is függetlenek legyenek egymástól. Ez onnan látszik, hogy az I_t indikátorváltozók összege pontosan akkor lesz binomiális eloszlású, ha azok függetlenek.

Ezt úgy is értelmezhetjük, hogy egy realizált VaR-on túli veszteséget nem követ nagyobb valószínűséggel még egy VaR-on túli veszteség, mint egy bármilyen más realizációt.

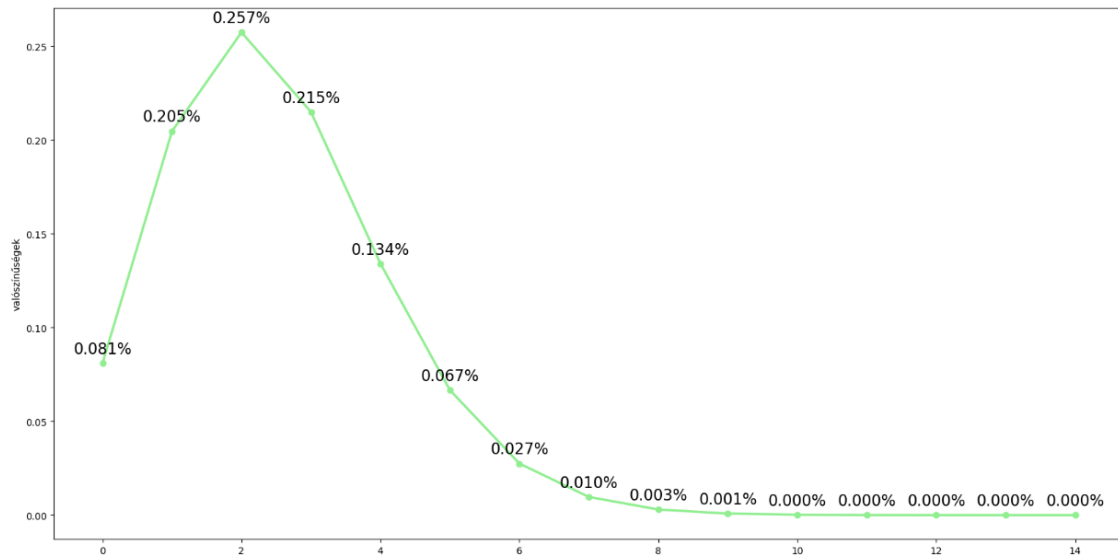
A szabályozói dokumentum egy példán keresztül illusztrálja ezt a módszertant és a hozzá tartozó statisztikai tulajdonságokat. Az [4] 1.táblázat 250 adatpont mellett mutatja meg a különböző percentilisek melletti elsőfajú hibát.

A modell helyes		
Breachek száma (250-ből)	Percentilis = 99%	
	tényleges	elsőfajú hiba
0	8,1 %	100,0 %
1	20,5 %	91,9 %
2	25,7 %	71,4 %
3	21,5 %	45,7 %
4	13,4 %	24,2 %
5	6,7 %	10,8 %
6	2,7 %	4,1 %
7	1,0 %	1,4 %
8	0,3 %	0,4 %
9	0,1 %	0,1 %
10	0,0 %	0,0 %
11	0,0 %	0,0 %

1. táblázat. A piaci kockázatok backtesting feltevései melletti valószínűségek és elsőfajú hibák, lásd: ([4], Table 1.)

Az 1. táblázat azt az esetet mutatja be, amit a pénzintézeteknek is alkalmazniuk kell, ha belső modell alapján kívánják a szavatolóátke-mennyiségüket meghatározni: egy év adatain végzett backtesting (250 adatpont) a veszteségeloszlás 99-es percentilisére (a $\text{VaR}_{99\%}(1 \text{ nap})$) azt feltételezve, hogy a megfigyelt breachek független statisztikai mintát alkotnak. A táblázat középső oszlopa az elméleti valószínűségeket mutatja: annak az esélye, hogy 250 független adatpont közül pontosan kétszer realizálunk a 99%-os percentilisének magasabb értéket, 25,7%. A harmadik oszlop értékei az elsőfajú hibát mutatják, tehát ha feltételezzük, hogy a modellünk jó, és elutasítjuk a modellt az adott mennyiségű breach mellett, akkor ekkora eséllyel fogunk tévedni. Így annak a valószínűsége, hogy 2 breach mellett egy jó modellt utasítunk el, 71,4%, és 0 breach mellett pedig 100%, hogy a modellünk jó volt.

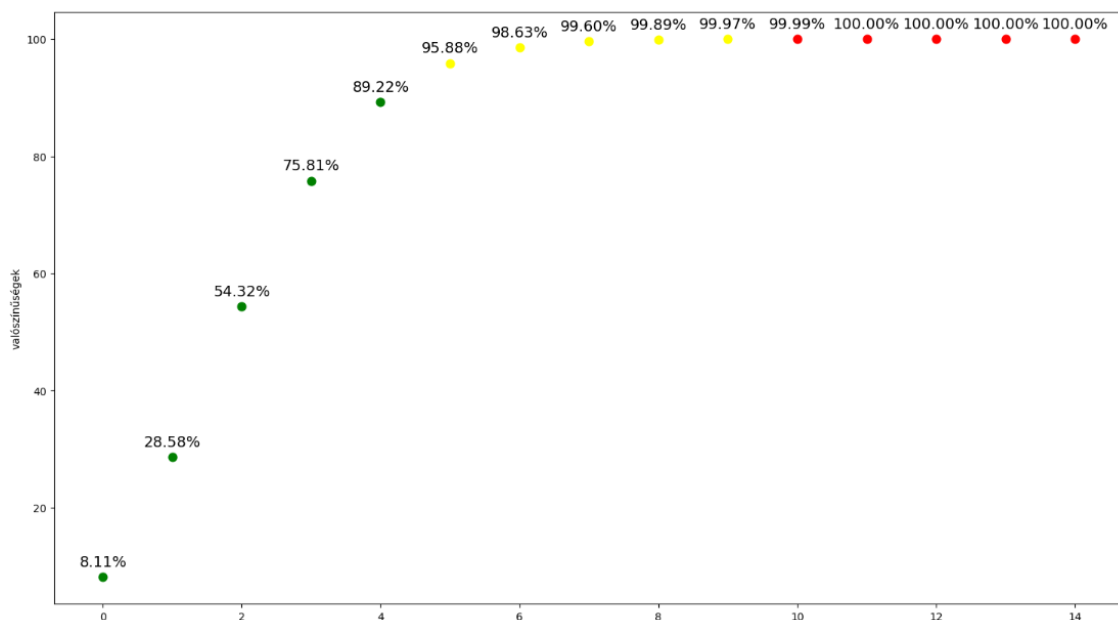
Ugyanezek az eredmények láthatóak a következő ábrán, ami a $(250; 1 - 0,99)$ paraméterű binomiális eloszlás adott ponthoz tartozó valószínűségeit mutatja meg:



1. ábra. A Binom(250; 0,01) eloszlás értékei néhány adatpont esetén

Ezáltal is láthatjuk, hogy az indikátor változóinkkal definiált breachek függetlennek vannak tekintve a szabályozás által is, hiszen ezek összege pontosan ebben az esetben ad vissza egy binomiális eloszlású véletlen változót.

A [4] 2. táblázata a 250 adatpont melletti három zónát mutatja meg, ami az 1. ábra kumulált valószínűségeiből kapható meg:



2. ábra. A Binom(250; 0,01) eloszlásfüggvénye és az értékhez tartozó zóna

A 2. ábra a (250; 0,01) paraméterű binomiális eloszlás eloszlásfüggvényét mutatja, és ezen értékek alapján van a három zóna kialakítva: például ha maximum 95% annak az

esélye elméletben, hogy egy adott mennyiségű breachnél nem realizálunk többet, akkor a modellünket jónak tekintjük, és ezáltal a zöld zónába soroljuk.

Abban az esetben, ha a backtesting alapján a modell a green zone-ba esik, nem kell megnövelni a pénzügyintézeteknek a szavatolótőkéjük mennyiségét, hiszen a modelltől nem származik olyan kockázat, ami nincs fedezve. Az azonban, hogy a modell green zone-ba esik, nem jelenti azt, hogy jól leírja a valóságot a veszteségeloszlásunk. Például egy olyan modell, ami 250 megfigyelésből egy $\text{VaR}_{99\%}$ -on túli veszteséget sem realizál, a green zone-ba esik, pedig ennek a valószínűsége csak 0,081. Egy ilyen helyzetben tehát valószínű, hogy a modellünk nem írja le jól a tényleges veszteségeloszlásunkat, azonban a szabályozás ezzel az esettel nem számol.

Ezen egyoldali megközelítésre ad alternatívát a Kupiec-teszt, aminek nullhipotézise a következő: $\alpha = \hat{\alpha} = X/N$, a teszt statisztikája pedig:

$$2 \log \left(\left(\frac{1 - \hat{\alpha}}{1 - \alpha} \right)^{N-X} \left(\frac{\hat{\alpha}}{\alpha} \right)^X \right),$$

ahol $\hat{\alpha}$ az empirikus excess ratio, α az elméleti excess ratio és X a megfigyelt excedek száma, N továbbra is a megfigyelések száma. A Kupiec teszthez hasonló kétoldali teszt Christoffersené is ([12]), azonban ezen dolgozat keretein belül csak az egyoldali megközelítéssel foglalkozunk.

Fontos megjegyezni, hogy az az 1996-os dokumentum ([4]), amiben a fent részletezett három zónás megközelítés került bemutatásra, a piaci kockázatok backtestingje. Míg a szavatolótőke-számítás alapja ezen a területen a 10 napos, 99%-os VaR, addig a backtestinget az egynapos, 99%-os VaR alapján számolják. A szabályozás előírja, hogy ezen adatokat legalább az elmúlt 12 hónap adataira kell tesztelni, így a 250 adatpontra kalibrált három zónás megközelítés összhangban van a piaci gyakorlattal.

Érdemes kitérni arra is, milyen hatással van a backtesting eredménye a pénzügyintézetek szavatolótőke-mennyiségére ([19]), hogy ezáltal is lássuk, milyen közvetlen hatással van ezen statisztikai eljárás a belső modelleket használó bankokra. A szabályozás által a piaci kockázatokra előírt szavatolótőke alapja a Bázeli III szerint a következő:

$$\text{Market Risk Charge} = (3 + x + y) \cdot M,$$

ahol az M a 10 napos, 99%-os VaR ($\text{VaR}_{0,99}(10 \text{ nap})$) és a stresszelt, 10 napos, 99%-os VaR ($\text{SVaR}_{0,99}(10 \text{ nap})$), azaz egy volatilisabb piacot feltételezve szimulált VaR) összege, az y a helyi szabályozói szervek által meghatározható tényező, az x pedig a backtesting eredmények alapján a következőképpen alakul:

Zóna	Exceedek száma 250-ból	x
Zöld	0 - 4	0,00
Sárga	5	0,40
	6	0,50
	7	0,65
	8	0,75
	9	0,85
Piros	10+	1,00

2. táblázat. A piaci kockázatok backtesting eredményei alapján meghatározott büntetőfaktor

A 2. táblázatban azt láthatjuk, hogy a korábban leírt háromzónás megközelítés alapján miként határozza meg a szabályozás a büntetőfaktorokat a piaci kockázatok esetén. Minél több VaR-on túli veszteséget realizált a bank, annál valószínűbb, hogy a veszteségeloszlásuk alapján félretett tőke mennyisége nem elégséges. A büntetőfaktorok lényege tehát az, hogy az esetlegesen rosszul számolt tőkemennyiséget megnövelje annyira, hogy az ne okozzon fedezetlen kockázatot. Hiszen a szabályozás célja az, hogy ne legyen olyan kockázata a pénzüintézeteknek, amire nincsenek valamilyen módon felkészülve.

3.2. Partnerkockázat

Részben a 2008-as nagy gazdasági világválság hatására megjelent a Bázeli III (A *global regulatory framework for more resilient banks and banking systems* ([6])), és vele a piaci és alapvető hitelkockázatok mellett a hitelkockázatok egy új típusa, a partnerkockázat (counterparty credit risk) is.

A piaci kockázatokéval szemben a partnerkockázat backtestingjére a szabályozás nem adott egy konkrét módszertant, csak irányelveket:

„Az IMM modellek backtestingje egy folyamatosan fejlődő folyamat, és egy olyan végleges módszertan, mint ami a piaci kockázatok területén is létezik, még nem került meghatározásra. Ennek a dokumentumnak nem célja konkrét módszerek vagy statisztikai tesztek előírása, sem pedig a bankok azon képességének korlátozása, hogy saját validációs technikáikat fejlesszék. Inkább az a célja, hogy felvázolja a módszertani megfontolások területeit és a meglévő backtesting keretrendszer esetleges javításait a bankokban, valamint megpróbálja tisztázni a fogalmakat és kifejezéseket.” ([7], 2.pont, saját fordítás)

A partnerkockázati modellek backtestingje általánosan a pénzintézet EPE (expected positive exposure), azaz a várható pozitív kitétség modelljének vizsgálatát jelenti, ami aggregálja a következőket:

- piaci faktorok szimulálása,
- derivatívák árazó modelljei, és
- collateral szimulálási modellek (collateral, haircut, margin).

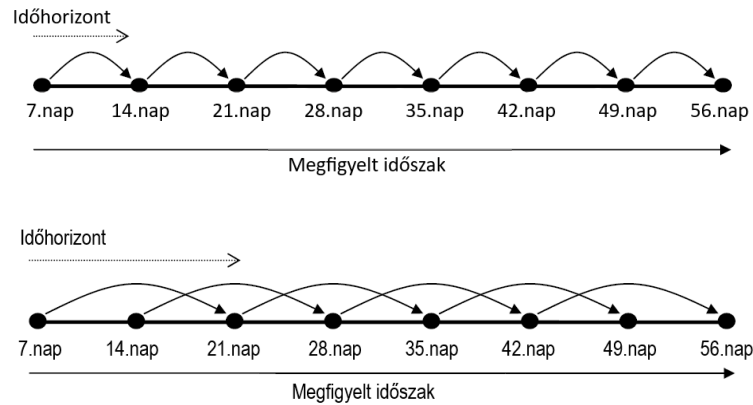
Ezen felül az is megnehezíti a modellezők helyzetét, hogy a fent felsoroltakat minden partnerre egyesével kell megvizsgálni, amihez egy kellően hosszú idősorra, változatlan portfólióösszetételre és konstans kereskedési díjakra lenne szükség. [7] alapján ezt a szinte lehetetlenül számításigényes folyamatot felválthatja a pénzintézet a saját, megfelelően kiválasztott hipotetikus ügyleteinek és portfólióinak backtestingje.

Érdemes megjegyezni, hogy míg a piaci kockázatok backtestingje az egy napos kockázatos értéket vizsgálja, addig a partnerkockázat esetén a szabályozás több időhorizontra is előírja az eljárás szükségességét (lásd: [7], 12-18.pont), sőt, kötelezi a pénzintézeteket az egy éves időhorizontra való szimulálás backtestelésére is ([7], 19.pont), hiszen az egy éves kitétségsprofil adja a Bázis III-as tőkeszámítás alapjául szolgáló EAD-t (Exposure-at-Default):

$$EAD = \alpha \cdot EEPE,$$

ahol az EEPE (Effective Expected Positive Exposure) az egyéves EPE (Expected Positive Exposure) profil monoton növekvő transzformáltjának idővel súlyozott átlaga ([24]).

Korábban láttuk, hogy a három zónás megközelítése a piaci kockázatoknak azon a feltevésen alapult, hogy a percentilisen túli megfigyelések előfordulásai függetlenek egymástól. Ez azonban hosszabb időhorizontok esetén már nem egy gyakorlatban használható feltevés az adatok mennyisége miatt (például 30 napos időhorizont mellett a piaci kockázatoknál használt 250 adatpontot több, mint 20 év alatt tudnánk összegyűjteni).



3. ábra. A különböző időtávok melletti sampling (saját szerkesztés, forrás: [7])

Ahogy azt a 3. ábrán is láthatjuk, egy rövidebb időhorizont (pl. egy hét) melletti független mintavétel továbbra is adhat nekünk elegendő adatot ahhoz, hogy a piaci kockázatoknál is használt binomiális tesztet a partnerkockázatok területén is alkalmazni tudjuk. Hosszabb időtáv mellett azonban már csak átfedő mintánk lehet, így a függetlenség feltétele nem teljesül.

A backtestingre használt statisztikai tesztek teljesítménye kisebb adathalmaz mellett tehát elsősorban a partnerkockázatok területének problémája.

A továbbiakban azon eseteket fogjuk csak vizsgálni, amelyeknél az adatok függetlenségét feltételezni tudjuk.

4. Binomiális teszt

Ebben a fejezetben a binomiális teszt kerül bemutatásra, ami az egyik legjobb backtesting módszer a fix paraméter melletti tesztek közül ([17]).

A binomiális teszt egy statisztikai hipotézisvizsgálat, amely nullhipotézise az, hogy a modell helyes, azaz az előrejelzett percentilis megfelel a valóságnak:

H_0 : a modell jó

H_1 : a modell nem jó

Egy adott szignifikanciaszint mellett (α) elutasítjuk a nullhipotézist a p percentilisére, ha a percentilisen túli megfigyelések száma (K) meghaladja a kritikus értéket (K^*):

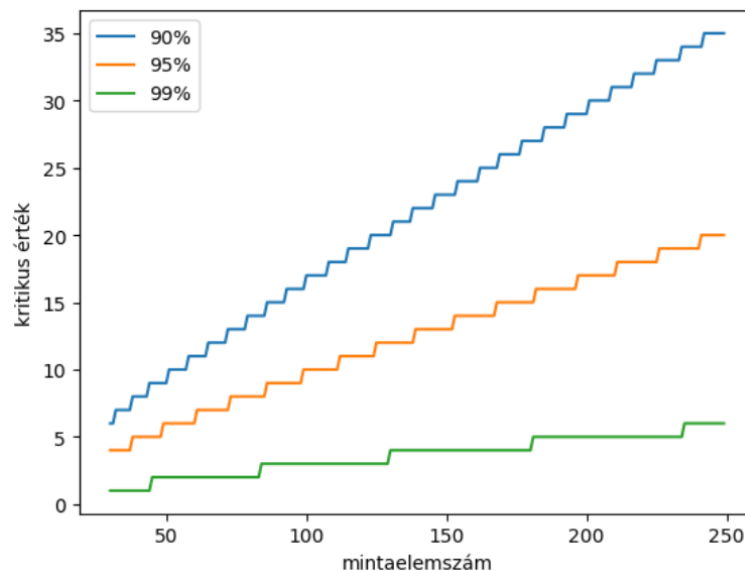
$$K^* = \min \left\{ K \mid \sum_{k=K}^N \binom{N}{k} p^k (1-p)^{N-k} \leq 1-\alpha \right\}$$

Ez a K^* érték eloszlásban tart a következőhöz a centrális határeloszlás tétele szerint:

$$K^* \xrightarrow{d} \Phi^{-1}(\alpha) \sqrt{Np(1-p)} + Np,$$

ahol a Φ a standard normális eloszlás eloszlásfüggvénye.

A kritikus értékek különböző mintanagyság esetén a következőképpen alakulnak 99%-os konfidenciaintervallum mellett:

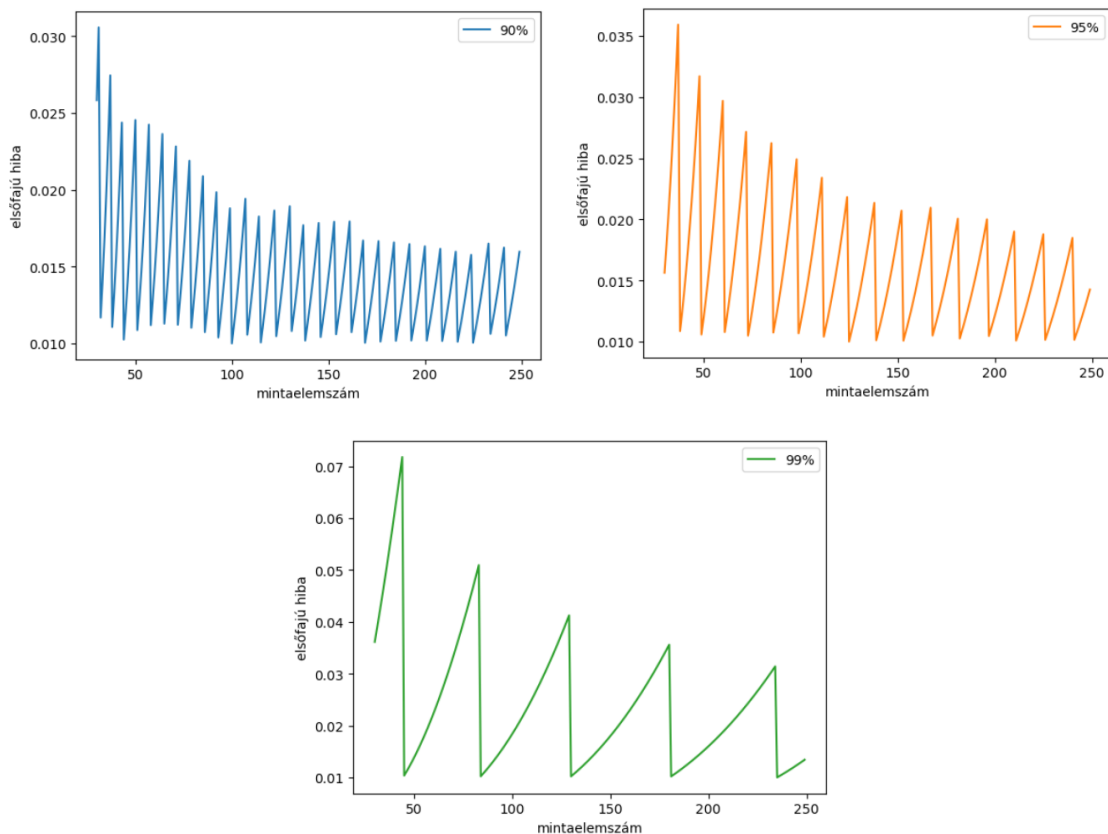


4. ábra. A különböző percentilisékhez tartozó kritikus értékek mintaelemszám függvényében

A 4. ábrán három felső percentilishez (90, 95, 99) tartozó kritikus értékeket láthatunk különböző mintaelemszámok esetén. Például, ha 100 adatpontunk van, akkor a 99-es percentilire végzett binomiális tesztet 3 vagy több percentilisen túli megfigyelés esetén utasítjuk el. Látható, hogy a binomiális eloszlás diszkrét volta miatt a kritikus érték a mintaelemszám függvényében egy monoton növekvő, lépcsős függvény.

Érdekes kiemelni, hogy míg a korábban látott háromzónás megközelítés egy rögzített adatmennyiség (250) mellett a szignifikanciaszint változtatásával érte el, hogy a 2. ábrához hasonlóan tudja értékelni a modelteljesítményt, addig ezen dolgozat a továbbiakban a szignifikanciaszint 99%-os szinten történő rögzítése mellett a különböző adatmennyiségekkel végzett tesztek teljesítményét vizsgálja.

A 4. ábrán is látható diszkrét mivolta a binomiális eloszlásnak az elsőfajú hibában is megjelenik:



5. ábra. Az eljárásból fakadó elsőfajú hiba mintaelemszám függvényében

Az 5. ábrán az elméleti elsőfajú hibákat láthatjuk különböző percentilisekre adatmennyiség függvényében. A statisztikai teszt kalibrálásához használt 1%-os érték, mint láthatjuk, nem mindig felel meg a valóságnak, ha kevés adat áll rendelkezésünkre. Ezen értékek kiszámításához a p percentilis és N elemszám esetén a $\text{Binom}(N; 0,01)$ eloszlás eloszlásfüggvényének értékét kellett meghatároznunk a kritikus értékének helyén:

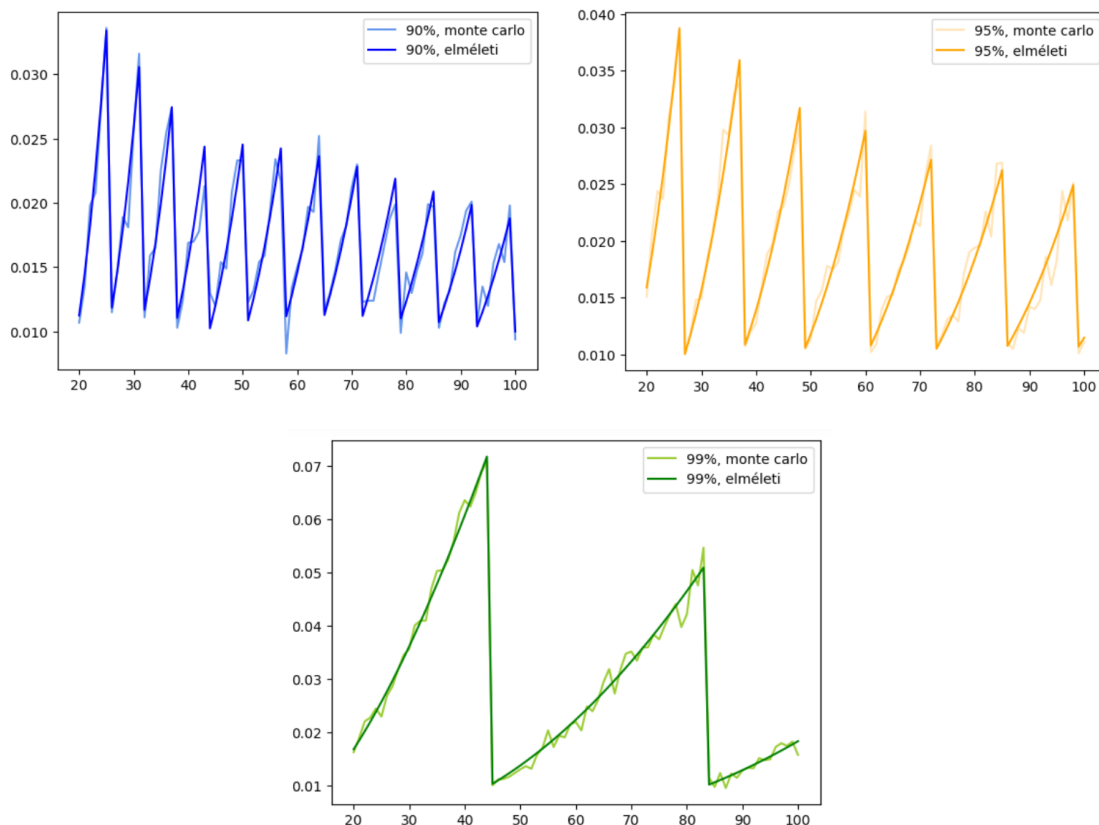
$$\sum_{k=0}^{K^*} \binom{N}{k} p^k (1-p)^{N-k},$$

majd egyből kivonva ezt az értéket kaptuk az elméleti elsőfajú hibát.

Megfigyelhetjük továbbá azt is, hogy minél kisebb a mintaelemszám és minél magasabb percentilist akarunk vizsgálni, annál magasabb lesz az eljárásból adódó elsőfajú hiba. Azt is láthatjuk, hogy a 4. és 5. ábra ugyanazon mintaelemértékeknél „ugrik”: minél magasabb az a mintaelemszám, amit még egy adott kritikus érték mellett tartunk számon, annál magasabb lesz az elsőfajú hiba. Ezzel együtt az adott kritikus értékhez tartozó legkisebb mintaelemszám mellett lesz a legkisebb elsőfajú hibánk.

Fontos megjegyezni, hogy az 5. ábra az elméleti elsőfajú hibát mutatja. A továbbiakban azonban az első- és másodfajú hibát Monte Carlo szimulációval fogjuk közelíteni.

Összehasonlításként végezzünk el egy Monte Carlo szimulációt az 5. ábra reprodukálására, 10000 iteráció mellett, $\mathcal{N}(0; 1)$ modellt feltételezve:



6. ábra. Az eljárásból fakadó és a szimulált elsőfajú hiba mintaelemszám függvényében

Az így kapott eredményeket összevetve a fent is bemutatott elméleti értékekkel láthatjuk, a szimulációhoz használt beállítások nagyon közel vannak az elméleti értékhez, tehát a

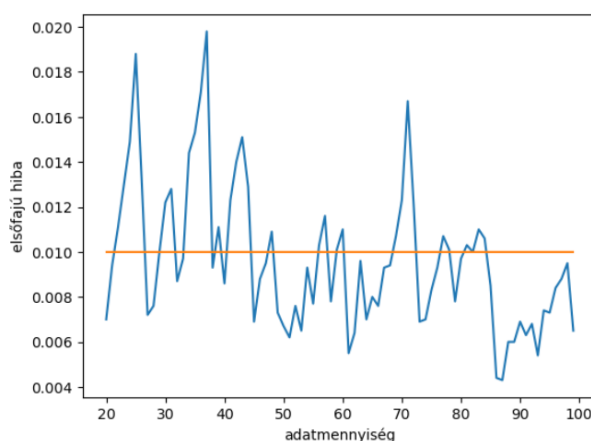
szimulációkat a továbbiakban így fogjuk végezni.

Az 5. ábra rámutat arra, hogy a jelenleg is piaci gyakorlatnak számító módszertan nem ad megbízható eredményeket alacsony mintaelemszám mellett. Azért is fontos egy olyan backtesting módszertant találni, ami alacsony mintaelemszám esetén is megbízhatóan kalibrálható elsőfajú hibát eredményez, mivel a szabályozás által a backtestinghez ajánlott minimális mintanagyság 20, ezáltal ezen félrevezető statisztikai eredmények a pénzügyi döntéshozatal részeivé válnak közvetett módon.

Fontos megjegyezni azonban, ahogy azt az ide vonatkozó szabályozói irodalom ([7]) 31. pontjában is kifejtik, hogy egy percentilis használata az eljárás során nem vezethet stabil eredményre abban a tekintetben sem, hogy a modell összteljesítménye jó-e, azonban több percentilis együttes használata mellett kevesebbszer kerül elutasításra a nullhipotézis:

„Jó példát mutatnak azok az IMM bankok, amik több percentilis figyelembevételével végzik a backtestingjüket. Továbbá, az egész eloszlás teljesítményének figyelembevétele növeli annak valószínűségét, hogy a modelleket ne utasítsák el egyetlen magas percentilisen alapuló gyenge teljesítmény miatt. Fontos, hogy az eloszlás tesztje ne egyetlen kockázati mérésre támaszkodjon, mind a szabályozási kitétség kiszámításához használt előrejelzési eloszlások integritásának értékelése érdekében, mind pedig azért, hogy elkerüljük a fals pozitív eredményeket, amelyek egyetlen extrém percentilis értékeléséből adódhatnak.” ([7], 31.pont, saját fordítás)

Ennek bemutatására tekintsük azt az eljárást, amikor a modellünk helyességét akkor utasítjuk el, ha a felső három percentilisből (90, 95, 99) legalább kettőt elutasítunk. Ekkor a 6. ábrához is használt beállítások mellett a következő elsőfajú hibákat kapjuk a korábban ismertetett szimulációs paraméterek mellett, adatmennyiség függvényében:



7. ábra. Az elsőfajú hiba nagysága a szimulált adatmennyiség függvényében, több percentilis együttes használata mellett

A 7. ábrát összehasonlítva a korábbi eredményekkel láthatjuk, hogy a vizsgált adatmennyiség (20 – 100) esetén több percentilis együttes használata a binomiális teszthez kisebb elsőfajú hibát eredményez, mint a percentilisek külön-külön történő backtestingje esetén tapasztaltak, tehát több percentilis együttes használata valóban csökkenti az elutasítások számát.

5. Bayesi statisztikán alapuló backtesting

5.1. Bevezetés, motiváció

Ezen fejezetben a bayesi statisztikán alapuló backtestinget vezetjük be a M. Zelvyte és M. Arnsdorf 2023-as tanulmánya, a *Bayesian backtesting for counterparty risk models* ([16]) alapján. Ahogy már azt korábban említettük, a piaci kockázatok területén ismert háromzónás megközelítés a következőn alapszik:

$$\sum_{k=0}^{\sum_{t=1}^N I_t} \binom{N}{k} p^k (1-p)^{N-k} = \alpha = \begin{cases} \text{red zone,} & \text{ha } \alpha \geq 0,9999, \\ \text{yellow zone,} & \text{ha } \alpha \geq 0,95, \\ \text{green zone} & \text{ha } \alpha < 0,95, \end{cases}$$

ahol N a mintanagyság, I_t a percentilisen túli megfigyelések indikátorváltozója, p pedig a vizsgált percentilis.

Az itt használt α ugyanazt az információt hordozza, mint a statisztikából ismert p -érték: a $0,95 \leq \alpha \leq 0,99$ feltétel megegyezik a $0,05 \geq p\text{-érték} \geq 0,01$ eredménnyel.

Ezen statisztikai mutató önmagában nem bizonyító erejű, ahogy azt [20] is kifejti, azonban ezen probléma eltörlül amellet a különbség mellett, hogy mi a backtesting eredeti célja, és mit jelent igazából a p -érték.

Míg a p -érték annak a valószínűsége, hogy az adott modellt tekintve nem fogunk a megfigyelt adatoknál extrémebbet realizálni a jövőben, ezt jelöljük a következőkben $p(\text{adat}|\text{modell})$ -el, addig a backtesting éppen azt a célt szolgálja, hogy a modellt ellenőrizzük a megfigyelt adatok függvényében, tehát a keresett valószínűség a $p(\text{modell}|\text{adat})$.

A hipotézisvizsgálatok jellemzője továbbá, hogy legfeljebb azt tudjuk mondani, hogy a modellünk helytelen, azt azonban nem, hogy az *miért*, vagy *mennyire* nem felel meg a valóságnak:

„A p -érték nem mondja meg, mennyire valószínű, hogy a modell helyes, vagy milyen mértékben van félrespecifikálva. A p -érték csak a nullhipotézis ellen szóló bizonyítékot jelezheti, de nem használható annak bizonyítására, hogy a modellünk helyes.” ([16], saját fordítás)

Ezen problémákat oldja fel a bayesi statisztika használata a frekventista megközelítéssel szemben. A keresett $p(\text{modell} | \text{adat})$ valószínűség felírható Bayes tétele alapján a következőképpen:

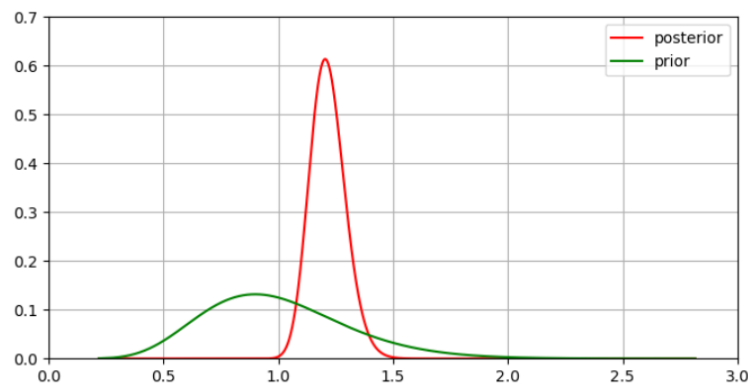
$$p(\text{modell} | \text{adat}) = \frac{p(\text{adat} | \text{modell}) \cdot p(\text{modell})}{p(\text{adat})}, \quad (2)$$

ahol a $p(\text{adat} \mid \text{modell})$ a likelihood, a $p(\text{modell})$ a prior, míg a $p(\text{modell} \mid \text{adat})$ a posterior eloszlás.

5.2. Példa a volatilitás félrespecifikálásra

Tekintsük a következő példát!

1. Példa. Adjon a modellünk $\mathcal{N}(0; 1)$ eloszlásból származó értékeket, míg érkezzenek a realizált értékek a $\mathcal{N}(0; 1,2)$ eloszlásból. Legyen a volatilitás prior eloszlása $\text{Gamma}(10; 10)$. Ekkor a likelihood függvény a normális eloszlás likelihood függvénye, a prior eloszlás a gamma. A következő ábrán azt láthatjuk, hogy az eredeti (prior) függvény hogyan alakult 50 realizált adatpont (kb egy év egyhetes horizonthoz tartozó adatmennyisége) felhasználásával egy olyan sűrűségfüggvénnyé, aminek a csúcsa körülbelül 1,2 felett van, ami pont a keresett volatilitás.



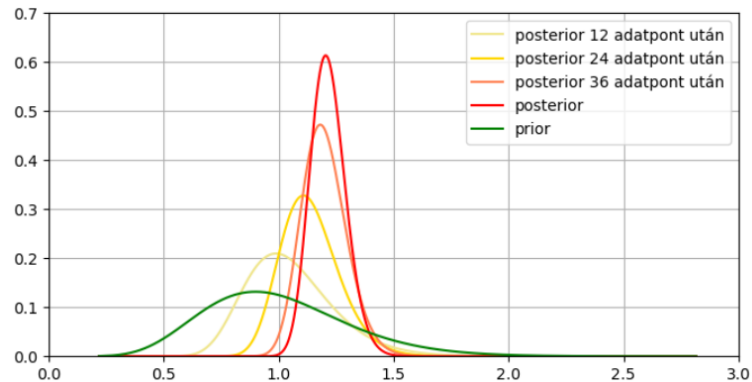
8. ábra. Példa a prior és posterior eloszlásokra

Tehát a bayesi megközelítést úgy is értelmezhetjük, hogy egy feltételezett eloszlást (a prior) a modell és a realizált adatok függvényével (likelihood) átsúlyozzuk, ezáltal kapunk egy olyan eloszlást (posterior), ami tükrözi a realizált adatainkat.

Ezen megközelítés tehát nem csak meghatározza a modell helyességét a posterior és a prior eloszlás „távolsága” alapján, hanem a posterior a backtestelt paraméterre vonatkozóan egy, a realizált adatokat jobban tükröző modellkalibrálást is elősegít.

Ezen megközelítés azt is lehetővé teszi továbbá, hogy a backtestelt idősort az idő függvényében is értelmezni tudjuk: míg a klasszikus, gyakoriságra támaszkodó tesztek nem tesznek különbséget az időben távolabbi és közelebbi adatpontok között, addig a fent említett átsúlyozást az adatok monoton növekvő halmazán is tudjuk értelmezni. Ezáltal a posterior eloszlások egy sorozatát is megkaphatjuk, ami segítheti a modellezőket abban, hogy jobban átlássák a háttér folyamatok változását.

Az 1. Példa alapján, ha nem az egész egy év adatait használtuk volna, hanem minden negyedévben elvégeztük volna a prior update-jét, akkor a következőt látnánk:

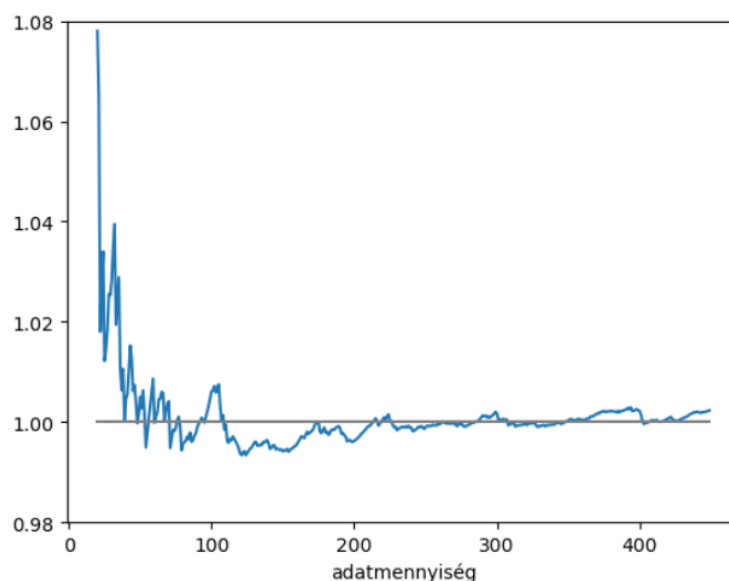


9. ábra. Példa a prior és posterior eloszlásokra

A 9. ábrán szépen látszik, ahogy egyre több adatunk van, egyre csúcsosabb a posterior, tehát egyre biztosabbak lehetünk a paraméterünk tényleges, a realizált adatok által mutatott értékében.

A posteriorok külön-külön adatrészekben vizsgálása hasznos lehet egy olyan többéves időszakban is például, amiben egy-két év válságos időszaka is benne van (pl. Covid időszak). A piaci folyamatok ekkor magasabb volatilitást mutattak, mint egy békésebb időszakban, így a válság lecsengése után a negatív eredményt adó backtesting hatására nem biztos, hogy érdemes a válságos időszaki paraméterekre változtatni a modelleket. A különböző időszakhoz tartozó adatok külön-külön vett posterior eloszlásából vissza tudunk következtetni arra, hogy a modellünk csak a válság miatt, vagy általánosan is rossz.

A 9. ábra alapján az lehet az intuíciónk, hogy a felhasznált adatmennyiség növekedésével a posterior eloszlásunk egyre csúcsosabb lesz, míg a végtelenben egy valószínűséget fog adni egy adott pontra. A következő ábra az 1. Példán bemutatott esetre vonatkozóan mutatja a posterior eloszlás várható értékének és a felhasznált adatok kumulált szórásának hányadosát adatmennyiség függvényében:



10. ábra. A posterior eloszlás várható értékének és a kumulált tapasztalati szórásnak a hányadosa adatmennyiség függvényében

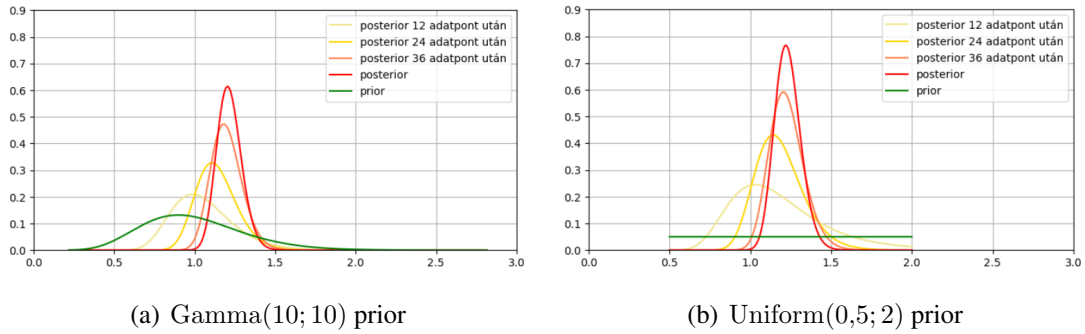
A 10. ábra egy N ponthoz hozzárendeli az első N pont tapasztalati szórásának és ezen N ponttal létrehozott posterior eloszlás várható értékének hányadosát. Láthatjuk, hogy viszonylag kevés (50) adat mellett is már egy-két százalék körül van ez a hiba, tehát a posterior eloszlás kevés adat mellett is megbízhatóan le tudja követni a tapasztalati szórást, és ezáltal a keresett paramétert is.

Megjegyezzük tovább, hogy a 9. ábra posteriorai nem feltétlenül az 1. példában meghatározott 1,2-höz kell, hogy tartssanak ilyen kevés adatpont esetén, hanem a realizációk tapasztalati szórásához, ami csak konvergál adatmennyiség függvényében 1,2-höz, azaz az elméleti értékhez.

5.3. Prior szerepe

Felmerülhet a kérdés, hogy milyen hatása van a prior választásának. Ennek vizsgálatára tekintsük a következő esetet: tegyük fel, hogy semmilyen korábbi tudásunk nincs a paraméterünk (volatilitás) eloszlásáról, és használjuk priornak a $[0,5; 2]$ intervallumon egyenletes eloszlást.

Ekkor az 1. Példán vizsgált idősoros megközelítés (10. ábra) a következőképpen módosul:



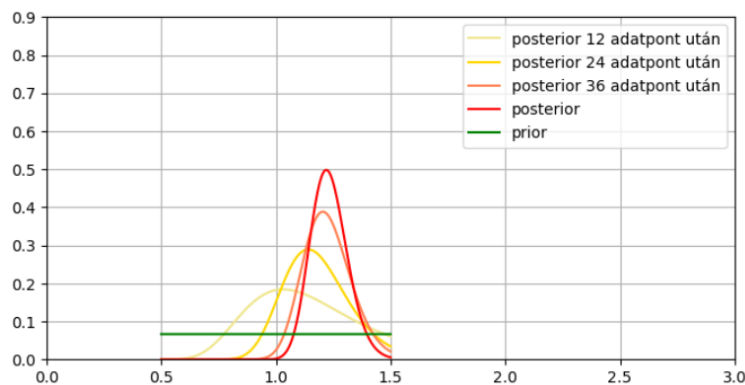
11. ábra. 1. Példa posterior függvényei Gamma(10; 10) és Uniform(0,5; 2) eloszlású priorról

A 11. ábrákon látott posteriorok várható értékei a következőként alakulnak:

adatmennyiség	12	24	36	50
tapasztalati szórás	0,911	1,184	1,261	1,237
11(a) ábra posteriorainak várható értéke	1,013	1,111	1,180	1,203
11(b) ábra posteriorainak várható értéke	1,164	1,188	1,229	1,235

Láthatjuk, hogy a 11(b). ábra posteriorai jobban (és mivel csúcsosabbak, ezért pontosabban is) követik le a tényleges, tapasztalati szórását az adatainknak. Ez azonban nem a prior sűrűségfüggvényének formájától, hanem a várható értékétől függött, hiszen míg az 50 realizált adat tapasztalati szórása 1,237, addig az Uniform(0,5; 2) eloszlás várható értéke 1,25.

Ha azonban egy olyan egyenletes eloszlású priort tekintünk, aminek a gamma eloszlással és a modellünkkel megegyező a várható értéke, akkor a következőképpen módosulnak az eredményeink:



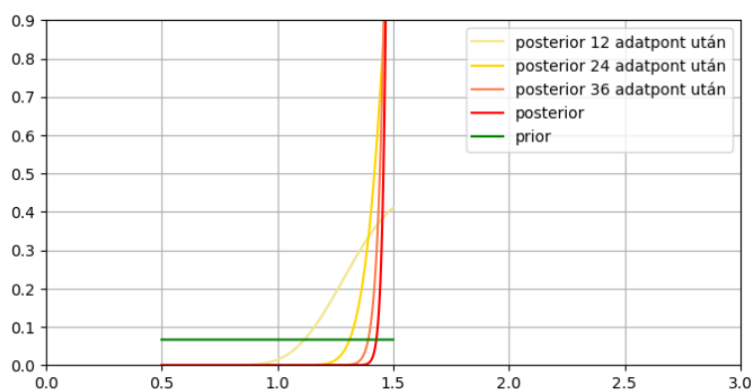
12. ábra. 1. Példa posterior függvényei Uniform(0,5; 1,5) eloszlású priorról

adatmennyiség	12	24	36	50
tapasztalati szórás	0,911	1,184	1,261	1,237
11(a) ábra posteriorainak várható értéke	1,013	1,111	1,180	1,203
12. ábra posteriorainak várható értéke	1,094	1,173	1,224	1,233
11(b) ábra posteriorainak várható értéke	1,164	1,188	1,229	1,235

Összehasonlítva a 11(a). és a 12. eredményeit azt láthatjuk, hogy érdemes a priort egyenletesnek megválasztani ahelyett, hogy a modellparaméterünk köré centrált sűrűségfüggvényűt választanánk.

Megjegyezzük továbbá, hogy a prior eloszlás megválasztása azért is nagyon fontos, mivel ha a realizált adatokat meghajtó keresett paraméterünk a prior eloszlás értékkészletében nincs benne, akkor fals eredményt ad a módszertan.

Ennek bemutatására tekintsük azt az esetet, amikor a prior eloszlás $\text{Uniform}(0,5; 1,5)$, azonban a realizált adatokat meghajtó folyamat volatilitása 1,9:



13. ábra. Bayesi backtesting, ha a keresett paraméter nincsen a prior értékkészletében

A 13. ábrán láthatjuk, hogy a posterioraink rásimulnak az intervallum szélére, ezzel fals, 1,5-ös paramétert adnak várható értékül a keresett 1,9 helyett, ezen hiba pedig csak abból ered, hogy a priornak választott $\text{Uniform}(0,5; 1,5)$ eloszlás nulla súlyt helyez a realizált adatok által mutatott paraméterértékre (1,9).

Ezek alapján azt a következtetést vonhatjuk le, hogy egyenletes eloszlást érdemesebb priornak használni, hiszen kevesebb adattal és pontosabban lehet vele a keresett paramétert meghatározni, azonban a korlátos értékkészlete miatt a szélsőségesen nagy vagy szélsőségesen kicsi paraméterértékeket nem tudja az eljárás lekezeln.

Érdemes lehet tehát a gyakorlatban kétlépcsősen használni a bayesi statisztikán alapuló eljárást: először priornak egyenletes eloszlást választani a legvalószínűbb intervallumra (például a $0,5; 2$), majd a szélső értékeket $(0,5; 2)$ adó eseteket külön megvizsgálni egy

másik prior mentén. Ezáltal a kevésbé szélsőséges eseteken kihasználhatjuk, hogy az egyenletes eloszlású prior mellett jobban leköveti a posterior a tapasztalati szórásnégyzetet, és a fals eredményeket is hatékonyan ki tudjuk szűrni.

Ez azonban tovább súlyosbítaná az eljárás legnagyobb hátrányát: a hatalmas számítás-igényt.

5.4. A prior és posterior cseréje

Ezen dolgozatban a modell és a realizáció eloszlása is homogén lesz az idő függvényében, függetlenül attól, hogy: „...*a modellek és a piacok nem statikusak, valamint a megfigyelések a piac állapotától függnnek a megfigyelt időszak elején. Ez azt jelenti, hogy a paraméterek és eloszlások különbözhetnek minden megfigyelt időpontban.*” ([16], 2.fejezet, saját fordítás).

Ugyanezen irodalom részletesen ki is tér arra, hogy különböző technikákkal miként lehet ezt az egynemű adathalmazt elérni, mi azonban ezt nem részletezzük.

Vezessük be formálisan a változóinkat a [16] jelöléseit használva. Legyen

- $\mathbf{y} = (y_i)_{i=1}^N$ a realizált homogén adatok idősora
- π a tesztelt paraméter prior eloszlása
- $(f_i)_{i=1}^N$ a szimulált és realizált adatok eloszlása
- $(\theta_i)_{i=1}^N$ azt a paramétert jelöli, amire kíváncsiak vagyunk, illetve ami meghatározza a vizsgált eloszlásokat.

Ezáltal az i -edik időpontbeli posterior eloszlás a következő:

$$P(\theta_i|y_i) = \frac{P(y_i|\theta_i) \cdot \pi(\theta_i)}{\int_{\theta_i} P(y_i|\theta_i) \cdot \pi(\theta_i) d\theta_i}. \quad (3)$$

Vegyük észre, hogy a nevezőben lévő mennyiség csak egy normalizáló tényező, ami azt garantálja, hogy a bal oldal valóban valószínűségi mérték legyen.

Ahhoz, hogy ezen prior-posterior szerepeket könnyen tudjuk felcserélni egymással („*today's posterior is tomorrow's prior*”), feltételezzük, hogy a modell félrespecifikálásának mértéke időben állandó, tehát $\theta_i = \theta$ minden i esetén. Úgy is gondolhatunk erre a mennyiségre, mint egy hosszú távú átlagos eltérésre, ami körül az egyes realizált θ_i -k ingadoznak, így az egyenlet a következőképpen módosul:

$$P(\theta|y_i) = \frac{P(y_i|\theta) \cdot \pi(\theta)}{\int_{\theta} P(y_i|\theta) \cdot \pi(\theta) d\theta}. \quad (4)$$

Mivel feltettük, hogy backtestinget csak függetlennek tekinthető adatokon végzünk, így a $P(y_i|\theta)$ mennyiségek függetlenek egymástól különböző y_i -k esetén. Ezáltal az y n elemű vektorra fennáll a következő: $P(\mathbf{y}|\theta) = \prod_{i=1}^n P(y_i|\theta)$, valamint az (2). egyenletben az y_i értékek helyett az y vektort is használhatjuk, azaz

$$P(\theta|\mathbf{y}) = \frac{P(\mathbf{y}|\theta) \cdot \pi(\theta)}{\int_{\theta} P(\mathbf{y}|\theta) \cdot \pi(\theta) d\theta}, \quad (5)$$

amit tekinthetünk a a posterior függvény definíciójának is.

Felmerülhet a kérdés, hogy az adatok függetlensége elegendő-e ahhoz, hogy biztosítsa, hogy a posterior függvénynek egy adott adathalmazzal történő update-elése megegyezik az adatok egyesével történő felhasználása után kapott posterior függvénnyel. Ezen kérdésre ad választ az alábbi levezetés:

Jelölje \mathbf{y}_n az (y_1, y_2, \dots, y_n) vektort. Ezen jelölések mellett

- $n = 1$ -re:

$$P(\theta|\mathbf{y}_1) = P(\theta|y_1) = \frac{P(y_1|\theta) \cdot \pi(\theta)}{\int_{\theta} P(y_1|\theta) \cdot \pi(\theta) d\theta}, \quad (6)$$

felhasználva a (4)-as egyenletet.

- $n = 2$ esetén

$$P(\theta|\mathbf{y}_2) = P(\theta|(y_1, y_2)) = \frac{P((y_1, y_2)|\theta) \cdot \pi(\theta)}{\int_{\theta} P((y_1, y_2)|\theta) \cdot \pi(\theta) d\theta}, \quad (7)$$

felhasználva, hogy az y értékeket függetlennek tekintjük egymástól, majd a számlálót és a nevezőt is leosztva a $\int_{\theta} P(y_1|\theta) \cdot \pi(\theta) d\theta$ mennyiséggel, felhasználva (5)-t, az egyenlet a következő alakra egyszerűsödik:

$$P(\theta|\mathbf{y}_2) = \frac{P(y_2|\theta) \cdot P(\mathbf{y}_1|\theta)}{\int_{\theta} P(y_2|\theta) \cdot P(\mathbf{y}_1|\theta) d\theta}. \quad (8)$$

- az $n + 1$ esetet tekintve a következőt kapjuk:

$$P(\theta|\mathbf{y}_{n+1}) = \frac{P(\mathbf{y}_{n+1}|\theta) \cdot \pi(\theta)}{\int_{\theta} P(\mathbf{y}_{n+1}|\theta) \cdot \pi(\theta) d\theta} = \frac{P(y_{n+1}|\theta) \cdot P(\mathbf{y}_n|\theta) \cdot \pi(\theta)}{\int_{\theta} P(y_{n+1}|\theta) \cdot P(\mathbf{y}_n|\theta) \cdot \pi(\theta) d\theta} \quad (9)$$

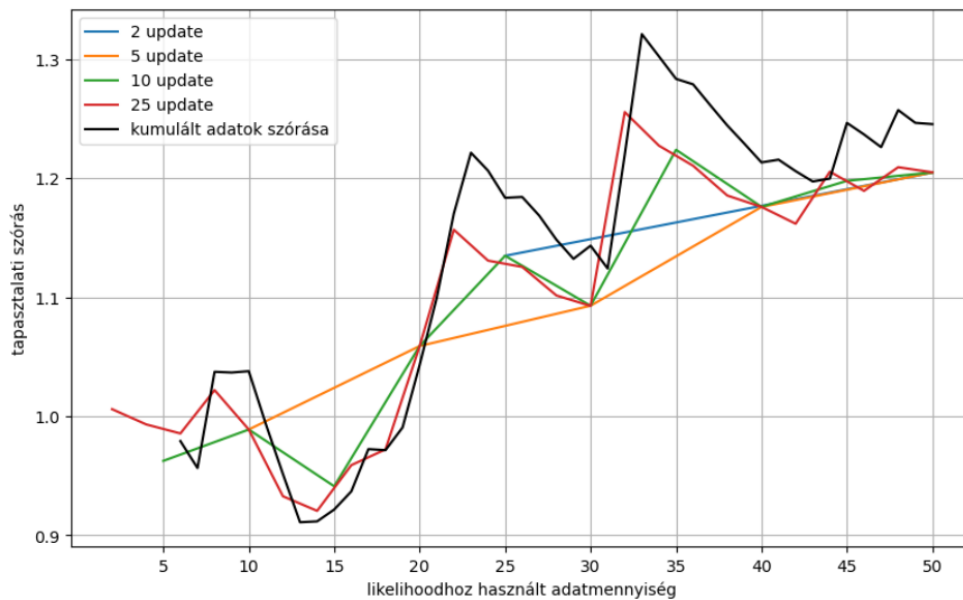
az $n = 2$ esethez hasonlóan, ha a számlálót és a nevezőt is leosztjuk a $\int_{\theta} P(\mathbf{y}_n|\theta) \cdot \pi(\theta) d\theta$ mennyiséggel, akkor a következő egyenletet kapjuk:

$$P(\theta|y_{n+1}) = \frac{P(y_{n+1}|\theta) \cdot P(\theta|y_n)}{\int_{\theta} P(y_{n+1}|\theta) \cdot P(\theta|y_n) d\theta}. \quad (10)$$

Megjegyezzük, hogy a fenti számítások szükséges feltétele, hogy a θ véges legyen, valamint hogy a sűrűségfüggvény paraméterek fölött vett integrálja nem nulla, véges érték legyen. A gyakorlatban ezen integrál kiértékelése teszi ki a számítási kapacitás nagyját még abban az esetben is, amikor a likelihood függvény felírható zárt alakban. Ilyen például az 1. Példában is használt normális eloszlás likelihood függvénye is. Vannak azonban olyan esetek, amikor a sűrűségfüggvény nem ismert, vagy nem írható fel zárt alakban, ekkor különböző közelítő eljárásokat lehet alkalmazni ([21], [22]).

A korábbi levezetés alapján azt állapíthatjuk meg, hogy nem számít, hogy az összes rendelkezésre álló adatot egyszerre felhasználva akarjuk frissíteni a prior eloszlást, vagy mindig egy-egy adatpontot felhasználva cseréljük ki a prior és posterior párokat, a végső posterior ugyanaz lesz.

Tekintsük ismét az 1. Példát! A realizált 50 adatpont tapasztalati szórását akarjuk a végső posterior eloszlással közelíteni, és ehhez kettő, öt, tíz, majd huszonöt lépésben cseréljük ki a prior-posterior párt. Ezután minden lépésnél összevetjük a kapott posterior eloszlás várható értékét a tényleges kumulált tapasztalati szórással:



14. ábra. A kumulált realizált adatok szórása és a különböző mintanagyságokon update-elt posteriorok várható értéke

A 14. ábra összehasonlítja a realizált adataink kumulált szórását (fekete) és a különböző gyakorisággal update-elt posteriorok várható értékét. A piros vonal például úgy készült,

hogy minden második megfigyelés után fel lett cserélve a prior és posterior, majd az eredményként kapott posterior eloszlás várható értékét ábrázoltuk. Hasonlóan, a kék vonal az első 25 megfigyeléssel update-elt posterior várható értékét mutatja a 25-ös ponthoz rendelve, majd a még fel nem használt adatokkal frissítve ezen posteriort kapjuk meg a végső eloszlásunkat, aminek várható értékét az 50-es ponthoz rendelt kék érték adja.

Láthatjuk, hogy a közös pontokon (például 10, 20, 50, 100) az eredmény ugyanaz lesz, azonban gyakoribb update-elés mellett jobban lekövetjük az adatok tapasztalati szórásának alakulását.

Ezen eredmények tehát azt mutatják, hogy a posteriort elegendő egyszer, az összes rendelkezésre álló adattal kiszámolni, hiszen a köztes prior-posterior váltások nem befolyásolják a végső eloszlást. Ez azért is nagyon szerencsés, mivel a likelihood kiszámítása sok időt vesz igénybe, azonban ha egyszerre használjuk fel az adatainkat, akkor ezt a mennyiséget elég csak egyszer meghatározni.

5.5. Hipotézisvizsgálat

Felmerül továbbá az a kérdés is, hogy miként lehet egy hipotézisvizsgálatot alapozni a bayesi megközelítéssel készített eljárásra.

A hipotézisvizsgálat nullhipotézise ugyanaz, mint a binomiális teszt esetén:

H_0 : a modell jó

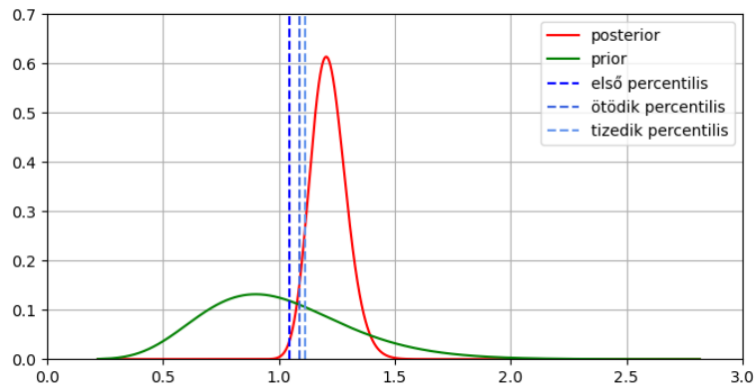
H_1 : a modell nem jó

Az α szignifikanciaszint mellett elutasítjuk a nullhipotézist, ha a korábban definiált posterior eloszlás (5) sűrűségfüggvényének α percentilise magasabb, mint a modellünk által feltételezett paraméter, tehát ha:

$$\inf \{x : P((\theta|\mathbf{y})(x) < 1 - \alpha)\} > \theta, \quad (11)$$

akkor elutasítjuk a nullhipotézist, miszerint a modellünk helyes.

Ha az 1. Példát tekintjük, akkor a modellünk $\mathcal{N}(0; 1)$, adatokat adott, míg a realizált értékek a $\mathcal{N}(0; 1,2)$ eloszlásból jöttek. Az alábbi ábrán láthatjuk a prior (zöld) és posterior (piros) eloszlások sűrűségfüggvényét, valamint a posterior első, ötödik és tizedik percentiliséit.



15. ábra. 1. Példa a posterior percentiliseivel

Ha az $\alpha = 99\%$ esetet tekintjük, azaz 99% -os megbízhatósággal végezzük a hipotézisvizsgálatunkat, akkor a modellünk által feltételezett szórást, azaz az 1 -et kellene összehasonlítanunk a posterior $1 - 0,99 = 0,01$, azaz első percentilisével, ami kerekítve $1,0474$. Mivel a posterior első percentilise magasabb, mint az eredeti modellparaméterünk, ezért ebben az esetben elutasítanánk a nullhipotézist, miszerint a modellünk jó.

Így az 1. Példában bemutatott esetben a bayesi statisztikán alapuló backtesting elutasítaná a helytelen modellt, valamint a posterior várható értéke a realizált adatok alapján $1,203$, azaz nemcsak azt tudtuk meg, hogy a modellünk nem helyes, de azt is, hogy az adott prior választása mellett a legvalószínűbb tényleges keresett paraméter (azaz a szórás) $1,203$, ami pont az, ami mellett a realizált adatokat is generáltuk.

6. Tesztek

Ebben a fejezetben különböző scenáriók mellett kerül összehasonlításra a 4. fejezetben bemutatott binomiális teszt és az 5. fejezetben bemutatott bayesi statisztikán alapuló backtesting módszertan próba erő szempontjából, különböző adatmennyiségek mellett.

A scenáriók a következők:

- i. a modellünk és a realizáció is normális eloszlású
- ii. a modellünk normális eloszlású, a realizáció student eloszlású

Mindkét scenárió az adatok volatilitását fogja vizsgálni, mivel a kockázatkezelés szempontjából ez a leglényegesebb, hiszen ez az előrejelző modellek által legkönnyebben kalibrálható paraméter.

A bayesi és a binomiális esetében is a statisztikai tesztek 99%-os szignifikancia mellett fogjuk végezni.

A binomiális tesztet a következőképpen fogjuk alkalmazni: tekintjük a 90, 95 és 99-es percentilisére végzett binomiális tesztek eredményét, és a nullhipotézisünket akkor fogjuk elutasítani, ha a három percentiliséből legalább kettőt elutasítottunk.

Ezen fejezetben bemutatott eredmények Monte Carlo szimulációval készültek. A kódok Python nyelven, a NumPy és a SciPy kiegészítő csomagok felhasználásával lettek megírva. A felhasznált véletlen számok a NumPy beépített random szám generátorával készültek, ezáltal a Mersenne Twister algoritmusra támaszkodtak².

A bayesi megközelítésnél használt két eloszlás, a gamma, mint prior, valamint a normális, mint a likelihoodhoz használt eloszlás sűrűségfüggvényét a NumPy csomag segítségével, beépített függvények mellőzésével számoltuk, hogy ezzel is lecsökkentsük a futási időket.

6.1. Első scenárió

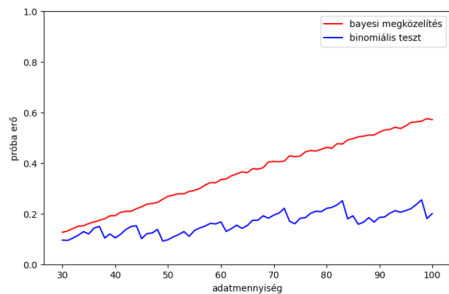
Ezen scenárió esetén a modellünk standard normális eloszlású, míg a realizáció szintén nulla várható értékű, de egynél nagyobb (1,2, 1,5, 1,8) szórású normális eloszlású.

A bayesi backtestinghez használt prior eloszlás a $\text{Gamma}(10; 10)$, míg a posteriorhoz használt likelihood függvény a normális eloszlás likelihood függvénye.

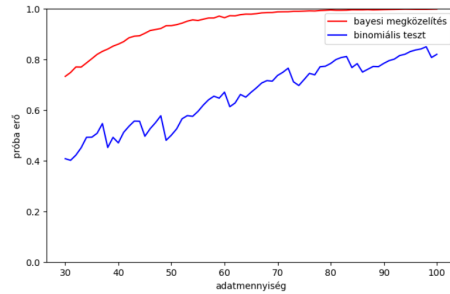
A cél a binomiális teszt és a bayesi statisztikán alapuló backtesting eljárás próba erejének összehasonlítása. A próba erő az elutasítás valószínűsége feltéve, hogy az alternatív hipotézis igaz.

²<https://numpy.org/doc/stable/reference/random/generator.html>

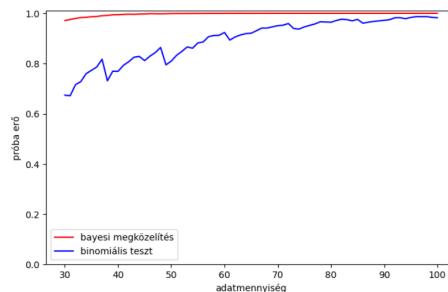
A következő ábrákon láthatjuk a három, magasabb volatilitás melletti elutasítási arányt, amit Monte Carlo szimulációval, 10000 iteráció mellett kaptunk. Mivel a realizációk volatilitása nem egyezik meg a modell volatilitásával, ezért ezen elutasítási arányok meg-
egyeznek a próba erejével, azaz az 1 – másodfajú hibával.



(a) volatilitás = 1,2



(b) volatilitás = 1,5



(c) volatilitás = 1,8

16. ábra. Próba erő összehasonlítása az első scenárió mellett

A 16. ábra mélyebb megértéséhez ragadjunk ki egy esetet és kövessük végig a két eljárást! Tekintsük az 1,5-ös volatilitás esetét, valamint a 60-as adatmennyiséget:

- i. a binomiális teszt elvégzéséhez generáltunk 60-szor 1000 standard normális eloszlásból származó adatpontot, amit a modellünk által szimulált eloszlásoknak tekintünk. Ezen eloszlások 90, 95 és 99-es percentiliséből képzett 60 hosszú idősorokat hasonlítjuk össze 60, 1,5-ös volatilitású normális eloszlásból származó véletlen számmal. Mindhárom percentilis esetén 99%-os megbízhatósággal binomiális tesztet végzünk, és ha a három percentilisből legalább kettő elutasítja a nullhipotézist, akkor a végső tesztünk is elutasítja a modell helyességét. Ezt az eljárást 10000 alkalommal megismételjük, és az elutasított esetek száma kb. 7000, innen jön a jobb felső ábra kék vonalának 60-hoz rendelt értéke
- ii. a bayesi backtestinghez szintén legeneráltunk 60 darab 1,5-ös volatilitású normális eloszlásból származó véletlen számot, majd ezen értékekkel update-eltük a $\text{Gamma}(10; 10)$ priort a 5. egyenlet alapján. A likelihood függvény a normális

eloszlás likelihood függvénye, hiszen a modellünk is ezen eloszlást feltételezi. A θ értékeit 10000 adat alapján határoztuk meg. Ha az így kapott posterior eloszlás első percentilise magasabb, mint egy, azaz a modellünk volatilitása, akkor elutasítottuk a nullhipotézist, miszerint a modellünk helyes. Ezt 10000-szer megismételve az elutasítások száma kb. 9200, és innen ered a 16(b) ábra 60-as pontjához tartozó piros érték.

Ezen szimulációt megismételve 30 – 100 adatponton és 1,2-es, valamint 1,8-as volatilitású normális eloszlásból származó realizációkkal kaptuk a 16. ábrát.

Láthatjuk, hogy ezen scenárió és ezen beállítások mellett minden esetben magasabb a próba ereje a bayesi megközelítésnek, mint a binomiális tesztnek. A 16. ábra alapján azt is megfigyelhetjük, hogy magasabb volatilitás mellett magasabb lesz a próba erő mindkét módszertan esetén, hiszen nagyobb valószínűséggel vesszük észre a modell helytelenségét.

A 16. ábra azt is megmutatja nekünk, hogy míg a bayesi megközelítés esetén a próba erő adatmennyiség függvényében egy viszonylag sima függvény, addig a binomiális teszt ereje adatmennyiség függvényében az elsőfajú hibához hasonlóan (5) egy hullamzó függvény. Ez alapján azt mondhatjuk, hogy a binomiális teszt sokkal érzékenyebb nemcsak a minta nagyságára, de a tényleges mennyiségére is.

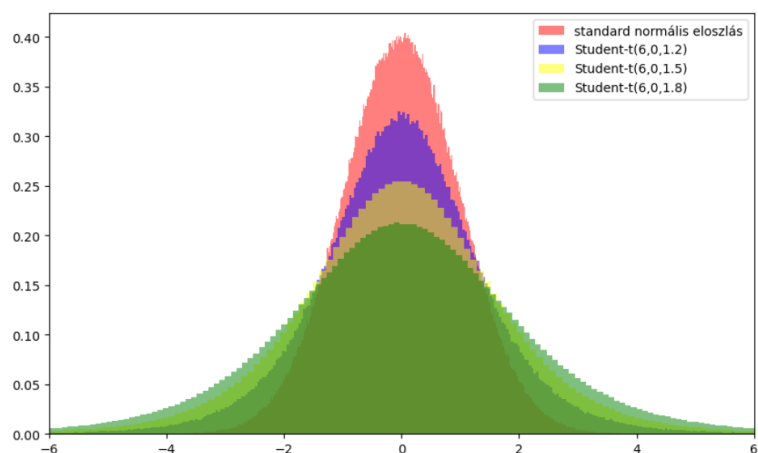
6.2. Második scenárió

A második vizsgált esetben a modellünk szintén standard normális eloszlású, és a likelihood függvény is ennek fényében a normális eloszlás likelihood függvénye.

Az első scenárióhoz hasonlóan a bayesi megközelítéshez használt prior a $\text{Gamma}(10; 10)$.

Ezen scenárió esetén azonban a realizált értékek 6 szabadságfokú Student-t eloszlást követnek nulla várható érték és rendre 1,2, 1,5 és 1,8 szórással.

A modell és a realizációk sűrűségfüggvénye ekkor a következőképpen alakulnak:

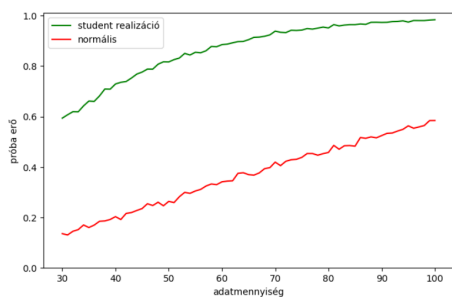


17. ábra. A második scenárió realizációinak hisztogramja

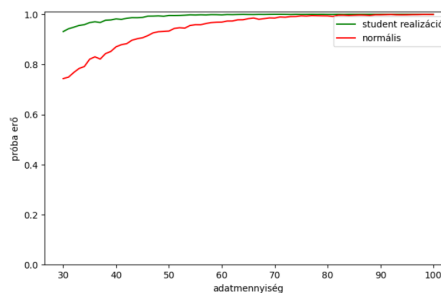
Láthatjuk, hogy a tesztekhez használt Student- t eloszlásoknak sokkal vastagabb a farka, mint a normális eloszlásnak. Így ezen scenárióval letesztelhetjük, hogy a bayesi statisztikán alapuló módszer alkalmas-e az ehhez hasonló, modellezési hibák észrevételére.

Arra a kérdésre is választ kaphatunk továbbá, hogy ha a modellünk nem normális eloszlású, akkor milyen hibát eredményez a normális eloszlás likelihood függvényének használata. Ez azért is hasznos, mivel a Student- t eloszlás likelihood függvényének kiszámítása számításigényesebb, mint a normálisé, valamint a szabadságfok helytelen megválasztása is befolyásolhatja az eredményt.

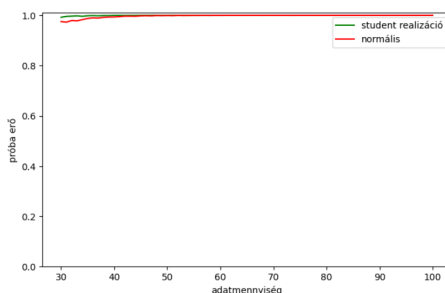
A következő ábrákon rendre az 1,2, 1,5, 1,8-as szórású normális eloszlású realizációk melletti próba erőket láthatjuk összehasonlítva az azonos szórású 6 szabadságfokú Student- t eloszlású realizációk melletti próba erővel, különböző adatmennyiségek mellett.



(a) volatilitás = 1,2



(b) volatilitás = 1,5



(c) volatilitás = 1,8

18. ábra. Próba erő összehasonlítása a második scenárió mellett

A 18. ábra ugyanúgy lett létrehozva, mint a 16. ábra bayesi backtesting eredményei. A 18. ábra piros grafikonja megegyezik a 16. ábra piros grafikonjával, ugyanúgy normális eloszlásból generáltuk a realizációkat hozzá. A 18. ábrán a zöld pontokat Student- t eloszlású realizációkkal generáltuk, azonban a bayesi backtestinghez használt többi paraméter ugyanaz, mint a korábban látottak.

Láthatjuk, hogy ha a modellünk normális eloszlású, és a bayesi backtesting során is a normális eloszlás likelihood-függvényét használjuk a Student- t eloszlásból jövő realizációk detektálásához, akkor is jól működik a módszer.

A 18. ábrán azt figyelhetjük meg, hogy a módszer sokkal megbízhatóbban működik, ha a realizációk Student- t eloszlásból származnak, mint ha azonos volatilitású normális eloszlásból származnának. Ennek az az oka, hogy a Student- t eloszlás egy vastagabb farkú eloszlás (ahogy azt a 17. ábrán is láthattuk), ezáltal a prior update-elése során több megfigyelés jön az eloszlás széleiből, mint a normális eloszlású realizációk esetén, ezáltal a tapasztalati szórás magasabb lesz kevesebb adat esetén is.

Ez egyrészt azért is lényeges, mivel a Student- t eloszlás likelihood-függvénye sokkal számításgényesebb, mint a normális eloszlásé, hiszen megjelenik benne a gamma-függvény, így futási idő szempontjából nem szerencsés választás. Másrészt modellezői szempontból az is fontos, hogy normális eloszlás feltételezése mellett is megbízható eredményeket ad-e a tesztünk, hiszen ez a feltételezés egy elterjedt piaci gyakorlat.

Megjegyezzük továbbá, hogy a 16. ábra binomiális teszt alapján meghatározott eredményeihez képest a bayesi backtesting eredményeinek legenerálása ötször annyi időt igényelt, ami jól mutatja, hogy mekkora az ugrás a futási időkből a két módszertan között. Ha a 18. ábrán a Student- t eloszlásból származó realizációkat a Student- t eloszlás likelihood függvényével értékeltük volna ki, akkor a binomiális teszthez képest körülbelül harmincszor több időt igényelt volna az ábrák elkészítése.

6.3. Eredmények összegzése

Ezen fejezetben a bayesi backtesting előnyeit és hátrányait fogjuk ismertetni az eddig látott eredmények alapján.

A módszer előnyei:

- próba erő szempontjából felülmúlja a binomiális tesztet,
- a próba erő az adatmennyiség kellően sima függvénye, így a binomiális eloszlásnál látott „ugrások” nem torzítják az eredményt,
- a posterior eloszlás ad a paraméter tényleges értékéről egy megfigyeléseken alapuló becslést,
- közvetlenül a keresett $p(\text{modelladat})$ értékkel foglalkozik a p -érték helyett.

A bayesi backtesting hátrányai:

- számításigényes,
- szükség van egy priorra, aminek a meghatározása egy egész piac szimulálásánál jelentős erőfeszítést igényel,
- a modellek komplexitása miatt a keresett paraméter meghatározása is bonyolult lehet.

Összességében a bayesi backtesting a partnerkockázatok területén lehet igazán hasznos, ahol a szabályozás nem határoz meg sem konkrét módszertant magára a backtestingre, sem a piaci kockázatokhoz hasonló büntetőfaktorokat. Az EPE modellek tesztelése során a volatilitás paraméterének meghatározása a posterior eloszlás segítségével segíthet egy esetleges tőkekiegészítés meghatározásában, valamint felgyorsíthatja a backtesting eredmények kiértékelését is.

7. Konklúzió, további kutatási kérdések

Ezen dolgozatban a jelenleg is használt elsődleges kockázati mérték, a Value-at-Risk (VaR), azaz a kockázatot érték backtestingjét vizsgáltuk meg klasszikus és bayesi statisztikai módszerekkel.

A 3. fejezetben bemutattuk a backtestinget, mint statisztika eljárást, valamint a szabályozói irodalom ide kapcsolódó részeit. Bemutatásra került a piaci kockázatok területén használt háromzónás módszertan, valamint a partnerkockázatok EPE (Expected Positive Exposure) modelljeinek backtestingje során felmerülő kihívások is.

A 4. fejezetben az egyik legjobb fix paraméter melletti backtesting módszertant vizsgáltuk meg, a binomiális tesztet. Kitértünk az alacsony mintaelemszám mellett fellépő magas elsőfajú hibára, valamint annak pontos okára és néhány tulajdonságára is.

A 5. fejezetben bevezettük a bayesi statisztikán alapuló backtestinget. Részletesen kitértünk a módszertan motivációjára, a p -értékkel kapcsolatos problémák bemutatására is. Ezen fejezetben egy konkrét példán keresztül mutattuk be a módszertan működését és szükséges bemeneti adatait abban az esetben, amikor a modell félrespecifikálása a volatilitás paraméteréből ered.

Megvizsgáltuk, hogy milyen gyorsan konvergál a posterior eloszlás várható értéke a realizált adatok tapasztalati szórásához, és arra a következtetésre jutottunk, hogy az eltérés nagyon hamar, pártíz adatpont után is kis hibával közelíti a keresett értékünket.

A bayesi backtestinghez a következőkre van szükség a klasszikus binomiális teszt bemeneti értékein felül: egy likelihood függvényre és egy prior eloszlásra. A 5. fejezetben kitértünk a prior szerepére is: megvizsgáltunk több lehetséges prior választást is, és azt lehet mondani, hogy az egyenletes eloszlás választásával gyorsabban konvergálunk a keresett paraméterhez, azonban az egyenletes eloszlás korlátos értékészlete miatt fals eredményt kaphatunk. Ennek feloldására javasoljuk a kétkörös backtestinget: először egyenletes eloszlást használva törekedjünk a valóságot minél pontosabban leíró paraméterek megtalálására, majd a prior eloszlás széleit visszaadó esetekben végezzük el újra az eljárást egy szélesebb értékészlettel rendelkező prior eloszlással.

Megvizsgáltuk továbbá, hogy a prior update-elésének gyakorisága milyen hatással van a végső posterior eloszlásra, és arra a következtetésre jutottunk, hogy az összes rendelkezésre álló adat felhasználásával készített posterior eloszlás pontosan megegyezik az időszakosan kicserélt prior-posterior párok után kapott végső posterior eloszlásával. Ezzel megmutattuk, hogy bár hasznos a posterior eloszlások várható értékének idősorát tekintve vizsgálni a mögöttes folyamatok dinamikáját, mégis elegendő az eljárást egyszer, minden

rendelkezésre álló adat felhasználásával elvégezni ahhoz, hogy megkapjuk a megfigyelt adatok által legjobban leírt keresett paramétert.

Az 5.5. alfejezetben bemutattuk, hogyan lehet a bayesi statisztikán alapuló backtestingre egy hipotézisvizsgálatot építeni.

A 6. fejezetben két scenáriót vizsgáltunk a volatilitás félrespecifikálására: az első esetben a modellünk és a megfigyelt adataink is normális eloszlásból származtak, míg a második esetben azt vizsgáltuk, hogy normális eloszlású modellt feltételezve hogyan alakulnak az eredmények, ha a megfigyelt adataink Student- t eloszlást követnek.

A 30 – 100 adاتمennyiség függvényében meghatároztuk a binomiális teszt és mindkét scenárió melletti bayesi backtesting próba erejét és azt tapasztaltuk, hogy a bayesi megközelítés minden esetben felülmúlta a binomiális tesztet.

A második scenáriót vizsgálva arra a következtetésre jutottunk, hogy az eljárás alkalmas nem csak a paraméterek, de a modellválasztás hibájának detektálására is egy olyan nagyon fontos esetben, amikor a megfigyelések eloszlása vastagfarkú, így a magasabb percentilisek backtestingjére kiemelten szükség van, hiszen a szélső értékeken tapasztalunk igazán nem fedezett kockázatot.

Összességében a bayesi backtesting felülmúlja a binomiális tesztet próba erő szempontjából, valamint lehetőséget ad a modellezőknek arra, hogy a posterior eloszlás segítségével jobban megértsék a realizált adatok mögött meghúzódó folyamatokat, ezáltal pontosabb modellkalibrálást tesz lehetővé. A bayesi backtesting nagy előnye továbbá, hogy pont a jó kérdésre ad választ: helyes-e a modellünk, ha bizonyítékként egy adott megfigyelt adathalmaz áll a rendelkezésünkre. Ezzel szemben, a klasszikus hipotézisvizsgálatban csak a megfigyelt adatok valószínűségét határozhatjuk meg egy adott modell specifikációjának helyessége esetén, és nem kapunk bizonyítékot magának a modellnek a helyességére vonatkozóan.

Az eljárás egyetlen általunk talált gyengeségét a számításigényessége adja, valamint, hogy a binomiális teszttel ellentétben egy plusz információra, a prior eloszlásra is szükségünk van, aminek a meghatározása szintén erőfeszítést igényel.

Hivatkozások

- [1] BCBS, *International Convergence of Capital Measurement and Capital Standards* (1988).
- [2] BCBS, *Amendment to the capital accord to incorporate market risks* (1996).
- [3] BCBS, *International Convergence of Capital Measurement and Capital Standards: a Revised Framework* (2004).
- [4] BCBS, *Supervisory framework for the use of 'backtesting' in conjunction with the internal models approach to market risk capital requirements* (2004).
- [5] BCBS, *Revisions to the Basel II market risk framework* (2009).
- [6] BCBS, *A global regulatory framework for more resilient banks and banking systems* (2010).
- [7] BCBS, *Sound practices for backtesting counterparty credit risk models* (2010).
- [8] BCBS, *Fundamental review of the trading book: A revised market risk framework* (2012).
- [9] J.P. Morgan, *Riskmetrics technical document* (1992).
- [10] B. Dömötör és JA. Miskó, *A piaci kockázat tőkekövetelményének szabályozása*, *Gazdaság és Pénzügy* 3 (3) (2016).
- [11] F. Anfuso, D. Karyampas, and A. Nawroth, *A Sound Basel III Compliant Framework for Backtesting Credit Exposure Models* (2013).
- [12] P. Christoffersen, *Evaluating Interval Forecasts*, *International Economic Review* **39** (1998), 841–862.
- [13] J. Berkowitz, *Testing Density Forecasts, With Applications to Risk Management*, *Journal of Business & Economic Statistics* **19** (2001), 465–474.
- [14] M. Wozniak and R. Slepaczuk, *LIGHT Benchmark - Comprehensive Backtesting Framework for Market Risk Models Comparison* (2023).
- [15] A. Shirazi, *Value at Risk (VaR) Backtesting Techniques and P-Value Risk Decomposition Analysis* (2014).
- [16] M. Zelyte and M. Arnsdorf, *Bayesian backtesting for counterparty risk models*, *Journal of Risk Model Validation* **17** (2023), no. 2.
- [17] D. Tasche, *Rating and Probability of Default Validation*. In: *Studies on the Validation of Internal Rating Systems*, BCBS **14** (2005).
- [18] C. Acerbi, *Coherent measures of risk in everyday market practice*, *Quantitative Finance* **7** (2007), 359–364.
- [19] I. Ruiz, *Backtesting Counterparty Risk: How Good is Your Model?*, *Journal of Credit Risk* **10** (2014).
- [20] R. L. Wasserstein and A. L. Schirm, *Moving to a World Beyond " $p < 0.05$ "*, *The American Statistician* **73** (2019), 1–19.
- [21] J. C. C. Chan and E. Eisenstat, *Marginal Likelihood Estimation with the Cross-Entropy Method*, *Econometric Reviews* **34** (2015), 256-285.
- [22] N. Friel and A. N. Pettitt, *Marginal Likelihood Estimation via Power Posteriors*, *Journal of the Royal Statistical Society Series B: Statistical Methodology* **70** (2008), 589-607.

- [23] A. H.-L. Lau, H.-S. Lau, and J.R. Wingender, *The Distribution of Stock Returns: New Evidence against the Stable Model*, *Journal of Business & Economic Statistics* **8** (1990), 217-223.
- [24] M. Sayah, *Counterparty Credit Risk in OTC Derivatives under Basel III*, *Journal of Mathematical Finance* **7** (2016), 1-38.