

Közzététel: 2021. március 9.

A tanulmány címe:

A mintavételezés hatása a pénzügyi adatok statisztikai tulajdonságaira és alkalmazása a kockázatkezelésben

Szerző:

VARGA JÁNOS ZOLTÁN, a Szegedi Tudományegyetem tanársegédje
E-mail: vjz@eco.u-szeged.hu

DOI: <https://doi.org/10.20311/stat2021.3.hu0233>

Az alábbi feltételek érvényesek minden, a Központi Statisztikai Hivatal (a továbbiakban: KSH) *Statisztikai Szemle* c. folyóiratában (a továbbiakban: Folyóirat) megjelenő tanulmányra. Felhasználó a tanulmány vagy annak részei felhasználásával egyidejűleg tudomásul veszi a jelen dokumentumban foglalt felhasználási feltételeket, és azokat magára nézve kötelezőnek fogadja el. Tudomásul veszi, hogy a jelen feltételek megszegéséből eredő valamennyi kárért felelősséggel tartozik.

1. A jogszabályi tartalom kivételével a tanulmányok a szerzői jogról szóló 1999. évi LXXVI. törvény (Sztj.) szerint szerzői műnek minősülnek. A szerzői jog jogosultja a KSH.
2. A KSH földrajzi és időbeli korlátozás nélküli, nem kizárólagos, nem átadható, térítésmentes felhasználási jogot biztosít a Felhasználó részére a tanulmány vonatkozásában.
3. A felhasználási jog keretében a Felhasználó jogosult a tanulmány:
 - a) oktatási és kutatási célú felhasználására (nyilvánosságra hozatalára és továbbítására a 4. pontban foglalt kivétellel) a Folyóirat és a szerző(k) feltüntetésével;
 - b) tartalmáról összefoglaló készítésére az írott és az elektronikus médiában a Folyóirat és a szerző(k) feltüntetésével;
 - c) részletének idézésére – az átvevő mű jellege és célja által indokolt terjedelemben és az eredetihez híven – a forrás, valamint az ott megjelölt szerző(k) megnevezésével.
4. A Felhasználó nem jogosult a tanulmány továbbértékesítésére, haszonszerzési célú felhasználására. Ez a korlátozás nem érinti a tanulmány felhasználásával előállított, de az Sztj. szerint önálló szerzői műnek minősülő mű ilyen célú felhasználását.
5. A tanulmány átdolgozása, újra publikálása tilos.
6. A 3. a)–c.) pontban foglaltak alapján a Folyóiratot és a szerző(ke)t az alábbiak szerint kell feltüntetni:

„*Forrás: Statisztikai Szemle* c. folyóirat 99. évfolyam 3. számában megjelent, *Varga János Zoltán* által írt, 'A mintavételezés hatása a pénzügyi adatok statisztikai tulajdonságaira és alkalmazása a kockázatkezelésben' című tanulmány (link csatolása)”

7. A Folyóiratban megjelenő tanulmányok kutatói véleményeket tükröznek, amelyek nem esnek szükségképpen egybe a KSH vagy a szerzők által képviselt intézmények hivatalos álláspontjával.

Varga János Zoltán

A mintavételezés hatása a pénzügyi adatok statisztikai tulajdonságaira és alkalmazása a kockázatkezelésben*

Effect of various sampling methods on the statistical properties of financial data and their application in risk management

VARGA JÁNOS ZOLTÁN, a Szegedi Tudományegyetem tanársegédje
E-mail: vjz@eco.u-szeged.hu

A pénzügyi adatok mintavételezése többféleképpen történhet. Leginkább az időalapú mintavétel a használatos. Lehetséges azonban olyan, ettől különböző módokon is végrehajtani az adatok gyűjtését és rögzítését, hogy azok jobban reprezentálják az adatgeneráló folyamatot, és a statisztikai következtetések levonására inkább alkalmasak. A szerző tranzakció-, volumen-, illetve összetékalapú adatsorokat állít elő egy NASDAQ-on jegyzett technológiai részvény napon belüli tranzakcióiból, illetve egy teljes év kereskedési adataiból, emellett a kockázattérték- (value at risk, VaR) modell segítségével ezen eljárások kockázat-előrejelzésben nyújtott teljesítményét vizsgálja. Az eredményei azt mutatják, hogy az alternatív mintavételezési módszerek esetében kedvezőbb statisztikai tulajdonságú (homoszkedasztikus, autokorreláció mentes, normális eloszlású) hozamsorokat lehet megfigyelni, továbbá az adott részvényenél a volumen- és összetékalapú adatsorok segítségével jobban működő kockázati modell kapható.

TÁRGYSZÓ: pénzügyi adatok, VaR, mintavételezés

Sampling of financial data can be performed in various ways, among which time-based sampling is the most common. Nevertheless, it is possible to collect and record data using other techniques, which may better represent the process of data generation, and make the statistical inference more accurate. In this study, samples are created using transaction data of a technological company's stock listed on NASDAQ, based on transactions (tick bar), volume (volume bar), and value (dollar bar) and are compared to regular time-based samples. The author's samples are used in a value-at-risk (VaR) model in order to assess their performance in risk management. The results

* A kutatást az EFOP-3.6.2-16-2017-00007 azonosító számú, „Az intelligens, fenntartható és inkluzív társadalom fejlesztésének aspektusai: társadalmi, technológiai, innovációs hálózatok a foglalkoztatásban és a digitális gazdaságban” című projekt támogatta. A projekt az Európai Unió támogatásával, az Európai Szociális Alap és Magyarország költségvetése társfinanszírozásában valósult meg.

suggest that the studied alternative sampling methods may provide financial data series which have favourable statistical properties. Moreover, the risk-management model performs better with value- and volume-based data series.

KEYWORD: financial data, VaR, sampling methods

A közgazdasági tanulmányok jelentős része empirikus jellegű, azaz valamilyen megfigyelés, adatgyűjtés, esetleg kísérlet bemutatását, értékelését és az azok alapján levont következtetéseket rögzítik a megfelelő és elfogadott statisztikai-ökonometriai eszközök alkalmazása mellett.¹ A kutatás kiindulási alapját a megfigyelés, mintavételezés során szerzett adatok jelentik, melyek minősége és tulajdonságai nagyban determinálják a kutatási következtetéseket is. Jelen tanulmány első-sorban a pénzügyi adatsorok különböző mintavételezési² eljárásokból fakadó statisztikai tulajdonságait próbálja vizsgálni. A pénzügyi adatokat (ár, kamat, árfolyam) alapvetően az idő függvényében rögzítik, mintavételezik: rendszeres időközönként (például naponta) mintavétel történik az adott instrumentum piaci tranzakcióiból, majd ezt követően a minta alapján rendelik hozzá az árat a kérdéses időszakhoz. Az így összeállított idősorok statisztikai tulajdonságai azonban nem minden esetben ideálisak: a hozamok eloszlása a normálistól jelentősen eltérő, vastagfarkú, illetve gyakran heteroszkedaszticitás és autokorreláció figyelhető meg (Kiss [2017]). A pénzügyi adatokat azonban nem kizárólag az idő függvényében lehet rögzíteni, hanem többek között a tranzakciók számának vagy az üzletkötések összértékének a függvényében is. Felmerül a kérdés, hogy alternatív mintavételezési eljárással kapott adatsorok vajon jobban reprezentálják-e a mögöttes adatgeneráló folyamatot, képesek-e jobb statisztikai következtetések levonásához hozzájárulni. A tanulmány célja kettős. Egyrészt bemutatja és végrehajtja a különböző pénzügyi mintavételezési technikákat, illetve összehasonlítja az így kapott adatsorok statisztikai tulajdonságait. Másrészt a módszerek egy lehetséges alkalmazását is ismerteti: egy standard VaR-modell³ segítségével az adatsorok kockázat-előrejelzésben, -mérésben nyújtott teljesítményét veti össze.

¹ „Manapság sok közgazdász inkább tekinthető alkalmazott statisztikusnak, aki adatokat elemez, rendszerez, és statisztikai eszközökkel információt nyer ki belőlük. A szükséges közgazdasági elmélet gyakran minimális, amire inkább szükség van, az a lehető legjobb statisztikai elemző képesség.” (Colander–Kupers [2014] 132. old. saját fordítás).

² Pénzügyi adatok esetén a teljes sokaság, azaz az összes tranzakció ismert. A mintavételezés célja a sokaság reprezentálása egy kezelhető méretű mintával.

³ A VaR a Bazel Bankszabályozás végrehajtását szabályozó, az Európai Parlament és a Tanács 575/2013/EU számú (Az Európai Unió Hivatalos Lapja [2013]) rendelete alapján is a bankok működése során felmerülő piaci kockázat mérésének egyik fontos módja.

A cikk felépítése a következő: az 1. és 2. fejezet a pénzügyi információszolgáltatás és a kereskedés technológiai evolúciójának főbb állomásait, a pénzügyi adatok konstrukcióját és legjellemzőbb tulajdonságait tekinti át röviden. A 3. és 4. fejezet az alternatív mintavételezési eljárásokat, illetve az adatokat és az alkalmazott módszertant mutatja be. Végül az 5. és 6. fejezet az eredmények tárgyalásával és a következtetések levonásával zárja a dolgozatot.

1. A pénzügyi adatok szolgáltatásának és kereskedésének technológiai fejlődése

A modern tőzsdék megalakulását követően a piaci résztvevők közötti kereskedelem a tőzsdei parketten személyesen történt, az adatokat papíralapon rögzítették. Az első fontos technológiai innováció a távíró megjelenése volt: ez lehetővé tette, hogy a kereskedőkhöz bekötött eszközök segítségével jussanak értékpapíripiaci adatokhoz. 1866-ban üzembe helyezték a transzatlanti távíró kábelt, amely az európai és észak-amerikai tőzsdék összeköttetését is megvalósította, nagyban hozzájárulva az időszakot jellemző intenzív pénzügyi globalizációhoz. A távírók sokáig Morse-kódot használtak, majd 1933-ban Németországban a posta (Reichspost) kiépítette az első telexhálózatot. Itt már a Morse-billentyűket írógép-billentyűzet váltotta fel, és hamarosan a főbb tőzsdék is telexen keresztül, képernyőre kivetített adatokat szolgáltatottak, akár közvetlenül az üzletkötők irodájába. A Quotron 1960-as bevezetésével a felhasználók tetszőleges értékpapíradatai is gyorsan és könnyen megismerhetővé váltak, nem csak azok, amelyek aktuálisan megjelentek a képernyőkön (a Quotronok kezdetben papírra nyomtatták a kért adatokat, de hamarosan, számítógép-alapú, kis képernyővel ellátott eszközökké váltak) (Saunders–Cornett [2021]).

Jelentős mérföldkő volt 1971-ben a NASDAQ megalapítása, amely az első teljesen elektronikus módon működő tőzsde volt. Kezdetben csak részvényárakra vonatkozó jegyzéseket szolgáltatott elektronikusán, később lehetővé vált a tranzakciók elektronikus módon történő végrehajtása is. Az 1990-es években, az internet terjedésével, megjelent az online kereskedelem. A technológiai változásoknak köszönhetően a számítógépek egyre fontosabb szerepet játszanak nemcsak az infrastruktúra fenntartásában és működtetésében, hanem az üzletkötésekben is. Az ún. programozott kereskedelem (program trading), amely legalább 15 értékpapír és minimum 1 millió dollár értékű tranzakció együttes végrehajtását jelenti, már az 1980-as években is a teljes kereskedési volumen látható és jelentős hányadát tette ki (Saunders–Cornett [2018]).⁴

⁴ <https://www.nyse.com/publicdocs/nyse/markets/nyse/PT122812.pdf>

A következő lépcsőfok az algoritmikus és a nagyfrekvenciás kereskedés (high-frequency trading, HFT) elterjedése volt. Előbbi arra vonatkozik, hogy a kereskedési stratégiát, részben vagy teljes egészében, számítógépes program segítségével, algoritmus valósítja meg. Utóbbi pedig olyan üzletkötéseket jelent, ahol az értékpapír tartási ideje rendkívül rövid, a másodperc törtrésze. Az automatikus vagy algoritmikus kereskedés kezdetben azt jelentette, hogy az adott tranzakciót feldarabolták kisebb részekre, és úgy hajtották végre. Például azért, mert az ügylet piachoz viszonyított mérete miatt alkalmas lehet az árak jelentős befolyásolására, ami túlzott költséghez/veszteséghez vezethet. Az elmúlt 20 évben az algoritmikus kereskedés kifinomult információfeldolgozó és -előrejelző modellek, illetve ezeken alapuló stratégiák alkalmazása révén fejlődött tovább. Mindezek együtt azt jelentik, hogy a nagy tőzsdéken az infrastruktúra változásának köszönhetően a kereskedés jellege alapvetően változott meg: a tranzakciók ma már elektronikus úton mennek végbe, a volumen jelentős része esetében pedig a kereskedési, befektetési döntések meghozatala előre beállított algoritmusok által történik (Jansen [2020]).⁵

2. A pénzügyi adatsorok tulajdonságai

A tőzsdén kereskedett pénzügyi vagy reáleszközök ára rendszertelenül és akár egy másodpercen belül is többször változhat, így a gyakorlatban az elemzés során a végrehajtott tranzakciókból valamilyen mintavételes eljárással nyerünk ki egy standardizált, kezelhető adatsort.

A mintavételezéshez az adatok összegyűjtése leggyakrabban az idő függvényében történik (az így kapott adatokat a szakirodalom time bars-nak hívja). A feladat, hogy valamilyen módon minden vizsgált időintervallumhoz hozzá kell rendelni egy árat, amely az adott időszakot reprezentálja. Ez számos módon történhet: ún. timestamp alkalmazásával, vagyis az adott időponthoz tartozó árat tekintjük az intervallumot reprezentáló árnak. Gyakran használatos az időszak volumennel súlyozott átlagos ára, az időszak nyitó- vagy záróára, esetleg az időszak legalacsonyabb vagy legmagasabb ára (Lopez de Prado [2018]).

Természetesnek tűnik az adatok idő függvényében történő mintavételezése, az eljárás azonban nem problémamentes. Gyakori feltevés, hogy a pénzügyi árak véletlen bolyongás folyamattal modellezhetők, azaz jelölje $Z(t)$ a t . időponthoz tartozó

⁵ A HFT teszi ki az Egyesült Államok részvénykereskedési volumenének több, mint felét, a tőzsdei határidős ügyletek esetében ez akár a 80 százalékot is elérheti (Miller–Shorter [2016]).

árat, ekkor a $Z(t+1) - Z(t)$ különbség független, normális eloszlású valószínűségi változónak tekinthető 0 várható értékkel. Azaz

$$Z(t+1) - Z(t) \sim N(0, 1).$$

Samuelson [1965] és *Mandelbrot* [1966] használták először a véletlen bolyongás helyett a martingálalt a folyamat leírására. A martingál kevesebb megszorítást tartalmaz, mint a véletlen bolyongás, a közgazdasági megfontolások alapján felépített modellek eredményeként inkább juthatunk ilyen folyamathoz, emiatt a szakirodalomban egyre inkább ennek a használata terjedt el. A martingál formálisan a következő: legyen $\{x_t\} (t \in [0, \infty))$ sztochasztikus folyamat. Ekkor $\{x_t\}$ martingál a Φ_t információhalmaz-sorozatra tekintettel, amennyiben $E[x_{t+1} | \Phi_t] = x_t$, azaz, ha $\{x_t\}$ martingál, akkor a rendelkezésre álló információ Φ_t alapján az X_{t+1} -re vonatkozó legjobb becslés x_t (*Leroy* [1989]).

A pénzügyi adatok eloszlására vonatkozóan a legegyszerűbb feltételezés, hogy az árak normális eloszlást követnek. Azonban ennek közgazdasági tartalma megkérdőjelezhető (a normális eloszlás szimmetrikus és alulról nem korlátozott), például a részvények ára nem lehet 0-nál kisebb. *Samuelson* éppen ilyen megfontolások miatt vezette be az árak logaritmusainak használatát (*Mackenzie* [2006]). A hozamok használata az árak helyett azzal az előnnyel jár, hogy a különböző árszínvonalú pénzügyi termékek lényegében normalizálhatók, és így összehasonlíthatóvá válnak.

A következő kézenfekvő feltevés tehát, hogy a hozamok eloszlása normális. Ám a normális eloszlás szimmetrikus, és alulról korlátlan volta itt is problémát jelent, egy hozam ugyanis nem lehet -1 -nél kisebb. Továbbá, amennyiben az egyperiódusú hozam eloszlása normális, akkor a több periódushoz tartozó hozam eloszlása nem lehet normális, tekintettel arra, hogy az az egyperiódusú hozamok szorzata.⁶ Ilyen megfontolások miatt a lognormális eloszlás használata tűnt kézenfekvőnek, vagyis egy olyan modell, ahol a hozamok logaritmusának eloszlása normális. A pénzügyi idősorok nem jellemezhetők teljeskörűen a lognormális eloszlás használatával. A historikus adatok ugyanis ferdeséget mutatnak, és a normálishoz képest nagyobb csúcossággal jellemezhetők, azaz a harmadik és negyedik momentum rendszerint nagyobb értékű (a normálishoz képest).

⁶ A tőkepiaci árfolyamok modellje (capital assets pricing model. CAPM) expliciten nem tartalmaz előírást a hozamok eloszlására vonatkozóan (tekintettel arra, hogy egy időszakos modellről van szó). A korai elméleti levezetésekben feltételként jelent meg a normális eloszlás, ez biztosította a pozitív meredekségű, konvex befektetői közömbösségi görbéket, amelyek szükségesek voltak a CAPM megoldásához. Később alternatív megoldásokat is találtak a közömbösségi görbék előállítására (lásd például *Berk* [1997]).

Formálisan, legyen x valószínűségi változó, μ várható értékkel és σ^2 szórással. A harmadik momentum ekkor (a ferdeségi mutató):

$$m_3 = E \left[\frac{(x - \mu)^3}{\sigma^3} \right]. \quad /1/$$

A negyedik momentum (csúcsosság) pedig a következő

$$m_4 = E \left[\frac{(x - \mu)^4}{\sigma^4} \right]. \quad /2/$$

A szimmetrikus eloszlások ferdesége 0, a normális eloszlás csúcsossága 3. A historikus hozam adatok rendszerint pozitív harmadik momentummal, és 3-nál nagyobb negyedik momentummal rendelkeznek (*Campbell–Lo–MacKinlay* [1997], *Kiss* [2017]).

Mandelbrot javasolta először ezen tulajdonságok modellezésére az ún. stabilis eloszlások használatát (*Mandelbrot* [1963]). Stabilis eloszlások azok, amelyek lineáris kombinációja is ugyanazt a stabilis eloszlást eredményezi. A stabilis eloszlások egy speciális esetének tekinthető a normális eloszlás, ezért Mandelbrot érvelése szerint a stabilis eloszlásokon alapuló pénzügyi modellek tekinthetők a korábbi modellek egyfajta természetes általánosításának is. A stabilis eloszlások képesek megragadni a pénzpiaci adatokon megfigyelhető vastagfarkú jelenséget, vagyis azt, hogy az extrém események előfordulásának valószínűsége jelentősen nagyobb a normális eloszláshoz képest. Ami azt illeti, éppenséggel ez a probléma a stabilis eloszlások alkalmazásával: olyannyira vastagfarkúak, hogy a varianciájuk nem véges (*Campbell–Lo–MacKinlay* [1997]). Ez azért probléma, mert a statisztikai-ökonometriai tesztek, módszerek közül sok a centrális határeloszlás tételén nyugszik. A centrális határeloszlás tétele kimondja, hogy a véletlen valószínűségi változók összege normális eloszlást követ, függetlenül az összegben szereplő változók eloszlásától. A tétel azonban bizonyos követelményeket támaszt a valószínűségi változókkal szemben, melyek egyike, hogy azok véges varianciával rendelkezzenek. Mivel a stabilis eloszlások ezt nem biztosítják (a normális eloszlás kivételével), ezért alkalmazásuk a statisztikai eszköztár tekintélyes részének használatát tenné lehetetlenné. Emiatt a stabilis eloszlások alkalmazása gyakorlatilag az előző évszázad hetvenes éve óta megszűnt (*Mackenzie* [2006]).

A pénzügyi idősorok a normálistól eltérő eloszlás tulajdonságai mellett gyakran autokorreláció és heteroszkedaszticitás jeleit mutatják. Idősoros adatok esetén a k -ad rendű autokorreláció definíciója:

$$\rho(k) = \frac{\text{Cov}(r_t, r_{t+k})}{\sqrt{\text{Var}(r_t)}\sqrt{\text{Var}(r_k)}}, \quad /3/$$

ahol a k arra utal, hogy az idősor k periódussal korábbi tagjával számítódik a korreláció. Vagyis autokorreláció léte esetén az idősor értékei korrelálnak a korábbi értékekkel. A heteroszkedaszticitás pedig arra utal, hogy az adatok volatilitása nem állandó, nem független az időtől. A heteroszkedaszticitás a különböző pénzügyi eszközök közötti korrelációra épülő kockázatkezelési stratégiák alkalmazását teszi problémássá (Kiss [2017]).

3. Alternatív módok a pénzügyi mintavételezésre

Miként a korábbiakban már említettük, a pénzügyi idősorok előállítására leggyakrabban az adatokból rendszeres időközönként végrehajtott mintavétellel történik. A tranzakciók a kereskedési nap során nem egyenletesen köttetnek meg: közvetlenül a nyitást követően valamelyest átlag alatti a volumen alakulása, záráshoz közeledve azonban gyakran inkább élénkebb az aktivitás. Ez azt jelenti, hogy a fix időközönként vett minta alulmintavételezi a forgalmasabb periódusokat és túlmintavételezi a lassabbakat. Emellett az automatizált kereskedés elterjedésével, a kereskedési algoritmusok és a nagyfrekvenciás kereskedés használatával a tranzakciók természetes ritmusa még inkább elszakad az időtől, jobban kapcsolódik a számítógépek processzorának ciklusidejéhez, az egyes események, illetve azok meghatározott sorozata jelenti a kronológia alapját, nem pedig az eltelt idő (Easley–López De Prado–O’Hara [2012]).

Mandelbrot mutatott rá talán először, hogy az árak esetén a véletlen bolyongás feltevése nem feltétlenül igaz, azaz az időben egymást követő árváltozások nem véletlen bolyongást követnek. Azonban az egymást követő tranzakcióknál ez inkább teljesülhet, tehát javaslatuk szerint a kereskedés természetesebb időmérési eszköze az egymást követő tranzakciók sorozata. Más szavakkal, egyszerűen lehet igaz, hogy az adott, fix periódusonként tekintett árváltozás Pareto-eloszlást és az árak változása adott számú tranzakcióként normális eloszlást követ (Mandelbrot–Taylor [1967],

Brada–Ernst–Tassel [1966]). Ezt követően *Clark* [1973] kereskedési adatok elemzése alapján amellett érvelt, hogy a tranzakciók helyett a volumenalapú mintavétel még kedvezőbb tulajdonságokkal rendelkezhet. Mandelbrot elismerően nyilatkozott *Clark* munkájáról, azonban a gyakorlati alkalmazást illetően mind a tranzakció-, mind a volumenalapú kereskedési óra esetén szkeptikus volt. Mandelbrot szerint ugyanis a kereskedési döntések továbbra is kronologikusan történtek, emiatt az egyéb módszerek alapján számított előrejelzések nem voltak hasznosíthatók (*Easley–López De Prado–O’Hara* [2012]). Ahogy korábban utaltunk rá, azóta ez a helyzet megváltozott, köszönhetően az algoritmikus kereskedésnek, a gépi tanulási módszerek alkalmazásának; a kereskedés jelentős része zajlik olyan módon, hogy a kronologikus időnek kisebb súlya van. Emiatt az ilyen technikát alkalmazó kereskedők alternatív mintavételezési eljárásokat használnak a hozam adatok számítására.

López De Prado [2018] a különböző, nem az idő függvényében végrehajtott mintavételezési eljárásokat részletezi, a következőkben ezeket mutatjuk be röviden.

Kézenfekvő lehetőség, *Mandelbrot–Taylor* [1967] eredeti javaslata szerint, ha meghatározott számú tranzakcióból történik a mintavétel valamilyen módon (például volumennel súlyozott átlag számításával), és az fogja reprezentálni az adatsorunk egy tagját. A szakirodalom ezt tick bars-nak hívja, mi a továbbiakban tranzakcióalapú módszerként fogunk hivatkozni erre. A módszer hátránya, hogy nem tesz különbséget a tranzakciók nagysága között, azaz ugyanúgy 1-nek tekinti az 1 és a 10 000 terméket magában foglaló üzletkötést is. Jelölje P_k a k -adik tranzakcióhoz tartozó árat, $\{TB_i\}$ a tranzakcióalapú adatsort, n a tranzakciók számát jelöli, amelyből a mintavétel történik, és legyen az ár záróár. Ekkor

$$TB_i = P_{i \times n}. \quad /4/$$

A tranzakcióalapú eljárás mellett *Clark* [1973] munkája alapján konstruálhatunk ún. volume bars-t (továbbiakban volumenalapú eljárás): ebben az esetben a mintavétel akkor történik, amikor meghatározott számú termék cserél gazdát az adott piacon. Például, az adatsor egy tagját képezheti 1 000 részvény értékesítése, illetve az ahhoz hozzárendelt ár, függetlenül attól, hogy az mennyi tranzakciót igényelt. *Jones–Kaul–Lipson* [1994] a pénzügyi volatilitás és a tranzakciós/volumen adatok hatását vizsgálva azt találták, hogy a tranzakciók önmagukban járulnak hozzá a hozamok volatilitásához, a tranzakció méretének nincs információ tartalma. Ezen a nyomon haladva tovább *Ané–German* [2000] vetették össze a két módszert empirikusan, és azt a következtetést vonták le, hogy a tranzakciós adatokból megalkotott hozam adatsor eloszlása jobban közelíti a normális eloszlást, mint a volumen alapján képzett. *Aldrichy–Heckenbachz–Laughlin* [2016] tőzsdei határidős ügyleteket elemezve arra jutottak, hogy a tranzakcióalapú adatok normális eloszlást követnek.

Jelölje $\{VB_i\}$ a volumenalapú adatsort, Vol_k a k -adik tranzakció volumenét ($k = 1, 2, \dots$), v pedig a volumen nagyságát, amelyből a mintavétel történik. Ekkor:

$$VB_i = P_{p+m-1}, \text{ ahol } m = \min \left\{ T \mid \sum_{j=p}^T Vol_j \geq v \right\}. \quad /5/$$

Végül a harmadik, az ún. dollar bars (továbbiakban összértékalapú) módszer esetén mintavétel akkor történik, amikor egy előre meghatározott értéket elér az adott eszköz tranzakcióinak összértéke, függetlenül attól, hogy az mennyi tranzakción keresztül valósult meg, illetve attól is, hogy hány termék cseréje jelent meg ezekben a tranzakciókban. Ebben az esetben például a mintavétel akkor történik, amikor a kereskedett összérték eléri a 10 millió dollárt. A hozzárendelt ár pedig záróár, vagyis az utolsó megkötött ügylethez tartozó ár, amivel a kívánt érték (10 millió) elérhetővé vált. *López de Prado* [2018] érvelése szerint az ilyen összértékalapú mintavételezéssel sokkal stabilabb átlagos napi megfigyelésszám érhető el, a többi eljárás keretében létrehozott adatsorok napi számai jobban ingadoznak. Jelölje $\{DB_i\}$ az összértékalapú adatsort, d az összértéket, amelyből a mintavétel történik:

$$DB_i = P_{q+r-1}, \text{ ahol } r = \min \left\{ S \mid \sum_{j=q}^S Vol_j P_j \geq d \right\}. \quad /6/$$

Az előbbieket foglalja össze az 1. táblázat.

1. táblázat

A tanulmányban tárgyalt pénzügyi mintavételezési eljárások összefoglalása
(Summary of the financial market sampling methods discussed in the study)

Mintavételezési eljárás	Mintavétel gyakorisága	Példa
Időalapú	Rendszeres időközönként	Óráként, naponta, havonta
Tranzakcióalapú	Meghatározott számú üzletkötésenként	100 tranzakció
Volumenalapú	Akkor, ha a gazdát cserélt instrumentumok kumulatív száma eléri egy adott értéket	1 000 részvény adásvétele
Összértékalapú	Akkor, ha az ügyletek kumulatív összértéke eléri egy adott összeget	100 000 dollár

4. Adatok, módszertan

Az alternatív pénzügyi mintavételezési lehetőségek bemutatását követően a következőkben azok megvalósítása és egymással történő összevetése következik. Először napon belüli tranzakciós adatokon végeztem el a transzformációkat, hogy tranzakció, volumen és összérték szerinti adatsorokat kapjak, majd ezek után az eljárást végrehajtottam egy év kereskedési napjain is. Ezt követően VaR-módszerrel megvizsgáltam, hogy a kockázat előrejelzésében, számszerűsítésében miképpen teljesítenek a különböző mintavételezéssel kapott adatok.

A napi adatsort a Microsoft részvény 2019. május 2. kereskedési napjának tranzakciós adataiból készítettem, a nap során összesen 33 100 tranzakció történt, a záróár 126,21 dollár volt részvényenként. A 2018. október 14. és 2019. október 14. közötti időszak során a napi átlagos kereskedési érték 28 991 824 dollár volt, a vizsgált napon ez 27 350 200 dollár, a napi hozam pedig $-1,03$ százalék. Jellemzően a napon belüli adatok nyilvános adtabázisokból nem érhetők el, a tanulmányhoz a *Refinitiv Eikonon* keresztül töltöttem le azokat.

Az éves adatsorok alapját a Microsoft 2019. május 11. és 2020. május 10. közötti adatai szolgáltatták. Az időszak végére esik a COVID-19 járvány elterjedése, amelynek negatív hatása megjelenik a részvény árfolyamában is. Ez azért előnyös, mert a piaci kockázat számszerűsítésére szolgáló VaR-modellek volatilis, hektikus piaci környezetben rendszerint rosszabbul teljesítenek, ezáltal pontosabb képet kaphatunk az alternatív mintavételezési eljárások esetleges előnyeiről.

A tranzakcióalapú adatsor létrehozása során az első kérdés, hogy hány tranzakció legyen a mintavételezés alapja. Erre vonatkozóan nincsenek egyértelmű szabályok, az összehasonlítás érdekében 50 tranzakciót tekintettem egy egységnek, mert ezáltal lényegében ugyanannyi elemből álló adatsort generáltam, mint az egyperces, hagyományos adatok. Az ár meghatározására az utolsó tranzakció árát használtam (az egyperces adatok záróárasak, azaz szintén az adott időszak utolsó árát használják). Hasonló elvet követtem a volumen és az összérték esetében is, azaz rendre a kereskedett mennyiség és a dollárban számolt tranzakciók egységét úgy választottam meg, hogy közelítőleg annyi megfigyelésből álljon az adatsor, mint a percalapú kronologikus. Ez azt jelenti, hogy a volumenalapú adatok esetén 7 000 részvény adásvételéből történik a mintavételezés záróár használatával, az összérték alapú esetén pedig a kumulatív módon 100 000 dollárt elérő adásvételekből, szintén záróár alkalmazásával. Tehát, például a volumenalapú módszernél, amikor a kereskedett mennyiség eléri a 7 000 darabot, az a megfigyelés egy elemét fogja képezni, a hozzárendelt ár pedig az utolsó tranzakció során alkalmazott ár lesz.

Fontos megjegyezni, hogy hosszabb idősoroknál ez egyáltalán nem szempont, nem szükséges arra törekedni, hogy a megfigyelések száma valamilyen kronologikus módon mintavételezett eljárással azonos elemszámot állítson elő. Jelen tanulmány-

ban pusztán a jobb összehasonlíthatóság érdekében alkottam meg az adatokat ily módon.

Az egy évet lefedő kereskedési adatok esetében a hagyományos idősort a napi záróárak szolgáltatják. A volumen- és összértékalapú mintavételezés során a szükséges volumen, illetve összérték egységét rendre 2,09 millió, illetve 315,9 millió dollárban határoztam meg, amelyet az óránkénti volumenadatokról számítottam. Ezzel azt értem el, hogy (hasonlóan a napon belüli adatokhoz) a megfigyelések száma nagyjából megegyezik a kereskedési napok számával (248 volumen- és összérték-megfigyelés, szemben a 249 kereskedési nappal). Ebben az esetben tranzakcióalapú adatsort nem számítottam, tekintettel arra, hogy egy kereskedési napon átlagosan kb. 40 ezer tranzakció valósul meg az adott részvény esetében, ezért számítási szempontból nem célszerű tranzakcióalapon mintát venni egyéves időtávon.

Az adatok statisztikai tulajdonságainak elemzése a következő módokon történt. Az eloszlás normalitásának vizsgálatára a Jarque–Bera-tesztet hajtottam végre. A nullhipotézis normális eloszlásra, az alternatív hipotézis ennek hiányára utal. Autoregresszív feltételes heteroszkedasztikus Lagrange-féle multiplikátor- (autoregressive conditional heteroskedasticity Lagrange multiplier, ARCH LM) tesztet alkalmaztam a homoszkedaszticitás ellenőrzésére, a H_0 ebben az esetben homoszkedaszticitást, a H_1 pedig heteroszkedaszticitást feltételez. Végül, autokorrelációt is vizsgáltam Ljung–Box-féle Q-statisztika segítségével, itt a nullhipotézis az autokorreláció hiányát jelenti.

Miután a kívánt adataink rendelkezésre állnak, a VaR által a kockázatkezelésben való alkalmazhatóságukat vizsgálom. A VaR-módszer segítségével, arra a kérdésre kaphatunk választ, hogy adott valószínűség mellett (jelölje ezt p) a következő n kereskedési napon, a piacok normális működését feltételezve, mekkora a várható maximális veszteség adott eszköz vagy portfólió tartása kapcsán (Jorion [2001]). A VaR-értéket szokás abszolút és relatív értékben is megadni (Kovács–Marsi [2018]). Utóbbi esetben az értelmezés a következő: $(1 - p) \cdot 100$ százalékgig biztosak vagyunk, hogy a következő n kereskedési napon a hozam nem lesz rosszabb, mint a VaR-érték. Jellemzően az n értéke 1 vagy 10, a p értéke pedig leggyakrabban 0,01. A volumenalapú mintavételezés esetében természetesen nem kereskedési napokra, hanem adott számú volumenegységre, összértékalapú adatsornál pedig az összértékegységre értelmezzük a VaR-t.

Feltéve, hogy a hozamok normális eloszlásúak, a várható érték 0, $p = 0,01$, $n = 1$, és σ jelöli a hozamok szórását, akkor a VaR számítása a következő formulával adható meg:

$$VaR = -2,33\sigma. \quad /7/$$

Mivel egy év minden kereskedési napjára elvégezzük a számítást, ezért a volatilitás alakulását is modellezni kell valamiképpen. Ennek végrehajtására

a RiskMetrics-modellt használtam, amely szerint a variancia előrejelzése a jelenlegi volatilitás és hozam (R_t) súlyozott átlaga:⁷

$$\sigma_{t+1}^2 = 0,94\sigma_t^2 + 0,06R_t^2, \quad /8/$$

a VaR számítása pedig a következőképpen alakul:

$$VaR_{t+1} = -2,33\sigma_{t+1}. \quad /9/$$

A VaR számítása, illetve a volatilitás modellezése az előbbieknél szofisztikáltabb módokon is végrehajtható. Ezen módszereket jellemzően az adatsorok kedvezőtlen statisztikai tulajdonságai hívták életre, azonban jelen tanulmány éppenséggel az alternatív mintavételezési eljárások lehetséges előnyeit vizsgálja, ezért a VaR számítása során célszerű az egyszerűbb, normális eloszlás feltételezésére épülő eszközöket alkalmazni.

A kapott VaR-adatsorok összehasonlításának egyik módja, ha megvizsgáljuk azon napok, volumenegységek, illetve összértékegységek számát, ahol a hozam rosszabb volt, mint az adott időszakra számított VaR.⁸ Jelölje ezt a számot x . A definíció alapján p valószínűséggel $R_t < VaR_t$, ezért 100 megfigyelésből nagyjából p esetben kell ennek bekövetkeznie. Ha $x > p$, akkor a kockázat-előrejelző modell alábecsüli a kockázatot.

Formálisan is vizsgálhatjuk a kérdést a következő módon: legyen

$$I_t = \begin{cases} 1, & \text{ha } R_t < VaR_t \\ 0 & \text{egyébként} \end{cases} \quad /10/$$

a találati függvény, továbbá legyen

$$X = \sum_{i=1}^n I_i. \quad /11/$$

Amennyiben a VaR-modell megfelelően működik, akkor X a binomiális véletlen változó egy realizációja, azaz

$$H_0 : X \sim B(n, p).$$

Adott szignifikanciaszint mellett ezek alapján meg tudjuk határozni, hogy a vizsgált minta X értékének milyen intervallumban kell lennie. Ez 5 százalékos szignifikanciaszinten, nagyjából 250 megfigyelés,⁹ és $p = 0,01$ esetén $[0, 5]$. Azaz,

⁷ A RiskMetrics volatilitás modellje lényegében GARCH(1,1)-nek tekinthető (Christoffersen [2012]).

⁸ Ez az ún. backtesting egyik módja (Holton [2003]).

⁹ A „nagyjából” szót az indokolja, hogy a tényleges adatsorainkban 248, illetve 247 megfigyelés van, ez a minimális eltérés azonban nem befolyásolja az intervallumokat.

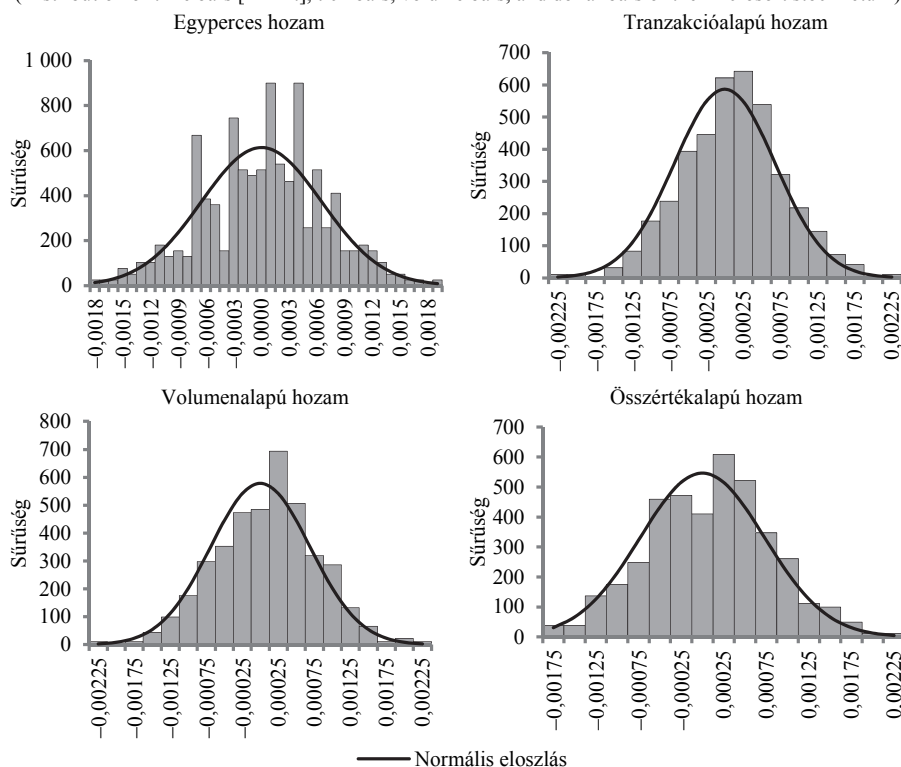
amennyiben $0 \leq x \leq 5$, nem tudjuk elutasítani a nullhipotézist, miszerint az adatso-
runk (a találati függvény) a megfelelő binomiális eloszlású sokaságból származó
minta. Ilyen feltételek mellett, amennyiben 5-nél többször marad el a tényleges ho-
zam a VaR előrejelzésétől, akkor a modellünk teljesítménye nem kielégítő.

A számítások, illetve statisztikai tesztek elvégzésére Matlab és EViews soft-
vereket használtam.

5. Eredmények

A 1. ábrán látható az egyperces hagyományos adatokból számított hozamok
hisztogramja és a megegyező szórással, átlaggal rendelkező normális eloszlás görbé-
je, illetve ugyanez a tranzakció-, volumen- és összértékalapú esetére is.

1. ábra. A Microsoft egyperces, tranzakció-, volumen- és összértékalapú hozamainak eloszlása
(Distribution of time bars [1 min.], tick bars, volume bars, and dollar bars of the Microsoft stock return)



Megjegyzés. Tranzakcióalapú hozam: 50 tranzakció = 1 egység; volumenalapú hozam: 7 000 részvény
értékesítése = 1 egység; összértékalapú hozam: 100 000 dolláros volumen = 1 egység.

Látható, hogy az eloszlás mindegyiknél normálisnak tekinthető, a Jarque–Bera- (JB) tesztstatisztikák az alternatív módszerek esetén kisebbek, vagyis jobban közelítenek a normálishoz. Homoszkedaszticitás alapvetően a tranzakció- és volumenalapú eljárás alkalmazásával jelenik meg az adatsorban, utóbbi p -értéke jóval meggyőzőbb (ARCH LM-teszt), míg autokorreláció egyik módszerrel vett mintára sem jellemző (Ljung–Box-teszt). Ezen adatok összefoglalóan a 2. táblázatban található. Az eredmények alapján a napon belüli adatok tekintetében a volumenalapon konstruált minta statisztikai tulajdonságai a legkedvezőbbek, de a másik három sem teljesít rosszul.

2. táblázat

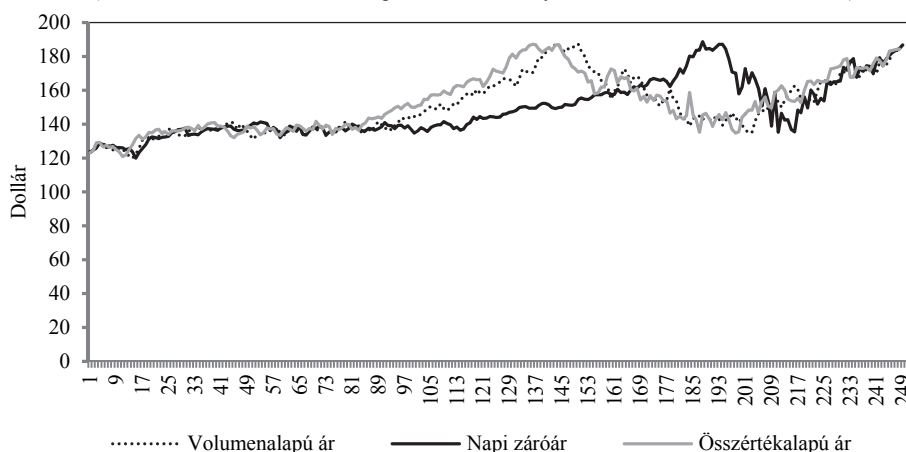
A napon belüli adatokból, különböző módokon mintavételezett részvényhozamok alapvető statisztikái
(Descriptive statistics of intraday data series obtained by various sampling methods)

Mutató	Egyperces hozam		Tranzakcióalapú hozam		Volumenalapú hozam		Összértékalapú hozam	
Átlag	0		0		0		0	
Medián	0		0		0		0	
Maximum	0,001809		0,002045		2,14E-03		0,002221	
Minimum	-0,001806		-0,002348		-0,002272		-0,001882	
Szórás	0,000652		0,000676		0,000689		0,000731	
Ferdeség	-0,08109		-0,05109		-0,02386		0,07748	
Csúcsosság	2,853577		3,176807		3,043865		2,826557	
Jarque–Bera-teszt; p -érték	0,774	0,679	0,671	0,715	0,064	0,969	0,726	0,696
ARCH-LM-teszt; p -érték	8,220	0,016	5,710	0,058	2,840	0,241	17,150	<0,010
Ljung–Box-féle Q-teszt; p -érték	1,722	0,423	3,217	0,210	3,459	0,177	2,224	0,329
Megfigyelések száma	389		386		363		322	

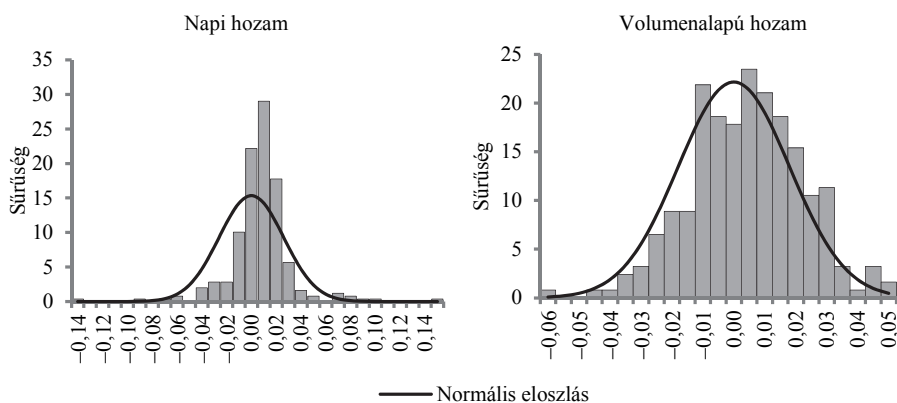
Megjegyzés. A nullhipotézisek a JB-, az ARCH-LM- és a Ljung–Box-teszt esetén rendre a normális eloszlás, a homoszkedaszticitás és az autokorreláció hiányát jelentik.

Éppen ezért érdekesebb az a kérdés, hogy vajon mi a történik, ha a bemutatott mintavételezési eljárásokat hosszabb időszakon, napi adatokkal vetjük össze. Az árfolyamok alakulását a 2. ábra szemlélteti. A kereskedési volumen nem állandó az adott időszak során, a COVID-19 járvány terjedésének hatására jelentősen megnövekszik, így az árfolyam visszaesése és a korrekció is elnyújtva jelenik meg a volumen- és összértékalapú árfolyamban. A hozamok eloszlásának hisztogramjai és a megfelelő normális eloszlás görbéi a 3. ábrán láthatók. Ez azt mutatja, hogy az adott időszak adatai alapján a volumenalapú mintavételezés segítségével a hozamok eloszlása a normálistól statisztikailag megkülönböztethetetlen.

2. ábra. A Microsoft részvény árfolyamának alakulása napi, volumenalapú és összértékalapú záróár alapján
(Evolution of the Microsoft stock price based on daily data, volume bars, and dollar bars)



3. ábra. A Microsoft napi és volumenalapú hozamainak eloszlása
(Distribution of time bars [daily] and volume bars of the Microsoft stock return)



Megjegyzés. Volumenalapú hozam: 2 090 000 dollár = 1 egység.

Az árfolyamokból számított hozamok statisztikáit a 3. táblázatban foglaltam össze. Látható, hogy a napi adatokból kalkulált hozam adatsora a JB-teszt alapján nem normális eloszlású, emellett autokorreláció és heteroszkedaszticitás jellemzi. A volumenalapú árfolyamhozam statisztikái ezzel szemben modellezési szempontból igen kedvező képet mutatnak: normális eloszlás, autokorreláció-mentes és homoszkedasztikus, vagyis a volumenalapú mintavételezéssel ismét egy valóban jobb tulajdonságokkal rendelkező, az ideálishoz közelítő adatsort sikerült generálni. Az összértékalapú adatok esetén vegyes az eredmény: a napi hozamokhoz képest kisebb csúcsosság tapasztalható, a JB-tesztstatisztika értéke egy nagyságrenddel

kisebb, azonban még így sem tekinthető normális eloszlásúnak. Ez a mintavételezési eljárás a heteroszkedaszticitást nem küszöbölte ki, azonban az autokorrelációt igen. Érdeemes még megjegyezni, hogy az alternatív adatsorok a napi hozamokhoz képest kisebb szórással rendelkeznek.

3. táblázat

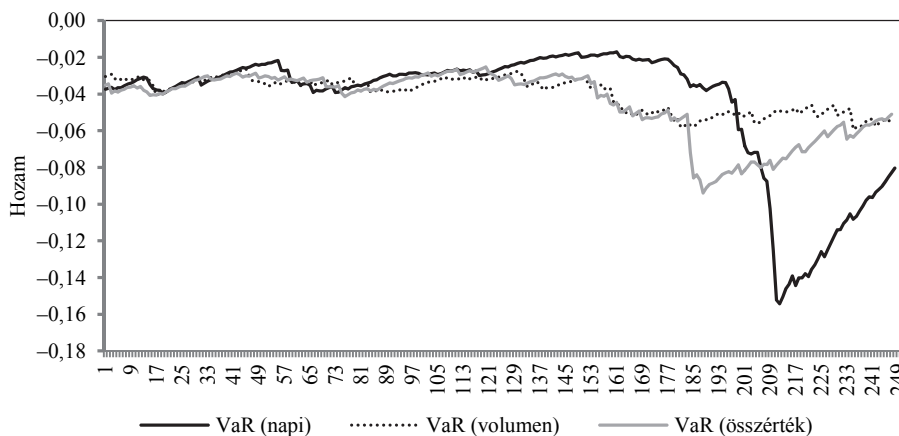
Egy év kereskedési adataiból, különböző módokon mintavételezett részvényhozamok alapvető statisztikái
(Descriptive statistics of 1-year data series obtained by various sampling methods)

Mutató	Napi hozam		Volumen alapú hozam		Összértékalapú hozam	
Átlag	0		0		0	
Medián	0		0		0	
Maximum	0,143094		0,049882		0,095846	
Minimum	-0,148353		-0,062841		-0,083622	
Szórás	0,025463		0,018165		0,021075	
Ferdeség	-0,040455		-0,139978		0,126396	
Csúcsosság	12,92568		3,25520		5,87846	
Jarque–Bera-teszt; p -érték	1 018,13	<0,01	1,48	0,48	85,93	<0,01
ARCH-LM-teszt; p -érték	47,81	<0,01	0,58	0,75	18,46	<0,01
Ljung–Box-féle Q-teszt; p -érték	55,40	<0,01	0,96	0,62	3,83	0,15
Megfigyelések száma	248		247		247	

Megjegyzés. A nullhipotézisek a JB-, az ARCH-LM- és a Ljung–Box-teszt esetén rendre normális eloszlás, homoszkedaszticitás, és autokorreláció hiányát jelentik.

Mindezek után érdemes megvizsgálni az eredmények kockázatkezelésben történő hasznosításának lehetőségét. A vizsgált időszakra kiszámított 1 napra, 1 volumen- és 1 összértékegységre számított VaR alakulása a 4. ábrán látható. Megfigyelhető, hogy mind a volumen-, mind az összértékalapon számított VaR sokkal kiegyensúlyozottabb, kevésbé volatilis. Az x értéke a napi VaR esetében 7, míg a volumen- és összértékalapon számított adatsoron 4, azaz mindkét módszer kevesebb hibát eredményezett, kevesebbszer becsülte alul a kockázatot (a 4. táblázatban mindez a VaR-túllépések száma sorban található).

4. ábra. A napi, a volumen-, illetve az összértékalapú adatsorok alapján számított VaR alakulása
(Evolution of VaR based on various data series [time bars, volume bars, and dollar bars])



A megfigyelések száma 247, és $n = 248$, továbbá $p = 0,01$, 5 százalékos szignifikanciaszinten az elfogadási intervallum $[0, 5]$. Tehát a volumen- és az összértékalapú VaR esetében 5 százalékos szignifikanciaszinten nem tudjuk elvetni a nullhipotézist, miszerint a minta megfelelő binomiális eloszlást követ, a napi VaR-nál azonban igen.

Az eltérések száma mellett fontos még azok átlagos nagysága is, ez szórásban kifejezve a napi, volumen- és összértékadatokon rendre 0,58, 0,35 és 0,34, tehát a napi hozamokon alkalmazott kockázat-előrejelző modell nemcsak gyakrabban, hanem nagyobb mértékben is alulbecsüli a kockázatot. Ez arra enged következtetni, hogy szemben a hagyományos napi statisztikák alapján számított előrejelzéssel, az alternatív mintavételezési eszközök jobban teljesítenek a kockázat-előrejelzésben adott részvény esetében. Mindezek összefoglalóan a 4. táblázatban láthatók.

4. táblázat

A VaR-előrejelzések teljesítménye
(Performance of VaR forecasts)

Megnevezés	Napi hozam	Volumenalapú hozam	Összértékalapú hozam
VaR-túllépések száma (x)	7	4*	4*
Túllépések átlagos nagysága (szórás)	0,5817	0,3428	0,3502
Megfigyelések száma	248	247	247

* Amennyiben x a $[0, 5]$ intervallumba esik, akkor 5 százalékos szignifikanciaszinten elfogadhatónak tartjuk a modell teljesítményét.

Az előzmények alapján – a volumen- és összértékalapú adatsoraink kedvezőbb statisztikai tulajdonságaik miatt (modellezési szempontból vizsgálva) – a VaR-megközelítés egy viszonylag egyszerű megvalósítása kedvezőtlen piaci környezetben is elfogadhatóan működik. Az eredményeink azt sugallják, hogy érdemes a mintavételezés módjára is figyelmet szentelni a pénzügyi adatok elemzése és a kockázatkezelés során.

6. Összefoglalás és következtetések

A megfigyelés és a kísérlet meghatározó szerepet tölt be a tudományos megismerésben, a megfigyelés, mintavételezés mikéntje azonban egyáltalán nem magától értetődő. A tanulmányban ezért röviden áttekintettem a pénzügyi idősorok mintavételezési kérdéseit. A kronologikus mintavétel, azaz a fix időintervallumhoz történő ár hozzárendelése statisztikai modellezési szempontból nem a legkedvezőbb. Már az 1960-as években Mandelbrot rámutatott arra, hogy a tranzakcióalapú adatsorok inkább a kívánatos normális eloszláshoz közelítenek. Nem sokkal később elterjedtek a volumen és összérték alapján generált minták is, hasonlóan előnyös statisztikai tulajdonságokkal. A vizsgálat során a Microsoft részvény napon belüli tranzakciós adataiból, illetve egy teljes év kereskedési adataiból is konstruáltam volumen-, illetve összértékalapú adatsorokat. Az eredmények alátámasztották a szakirodalom megállapításait, vagyis valóban jobb tulajdonságokkal leírható hozameloszlásokat lehet megfigyelni az újabb módszerek esetében, ez különösen a volumenalapúra volt igaz. Ezt követően egy egyszerű VaR-modell alkalmazásával teszteltem, hogy az éves adatsorok miként teljesítenek a kockázat erőjelzésében. Az eredmények alapján az alternatív mintavételezés segítségével jobban működő kockázatmodellt kaptam. A gépi tanulási eljárások pénzügyi használatával, az algoritmikus kereskedés elterjedésével, a nagyfrekvenciás kereskedők egyre fontosabb szerepével párhuzamosan a kockázatkezelésben is érdemes megvizsgálni az alternatív mintavételezési technikákon alapuló metódusok teljesítményét.

Irodalom

ALDRICHY, E. M. – HECKENBACHZ, I. – LAUGHLIN, G. [2016]: A compound duration model for high-frequency asset returns. *Journal of Empirical Finance*. Vol. 39. December. pp. 105–128. <http://dx.doi.org/10.1016/j.jempfin.2016.10.003>

- ANÉ, T. – GERMAN, H. [2000]: Order flow, transaction clock, and normality of asset returns. *The Journal of Finance*. Vol. 55. No. 5. pp. 2259–2284.
- AZ EURÓPAI UNIÓ HIVATALOS LAPJA [2013]: *Az Európai Parlament és a Tanács 575/2013/EU rendelete (2013. június 26.) a hitelintézetekre és befektetési vállalkozásokra vonatkozó prudenciális követelményekről és a 648/2012/EU rendelet módosításáról (EGT-vonatkozású szöveg)*. L 176. Június 27. 1–337. old.
- BRADA, J. C. – ERNST, H. – TASSEL, J. V. [1966]: The distribution of stock price differences: Gaussian after all? *Operations Research*. Vol. 14. No. 2. pp. 334–340.
- BERK, J. B. [1997]: Necessary conditions for the CAPM. *Journal of Economic Theory*. Vol. 73. Issue 1. pp. 245–257. <http://dx.doi.org/10.1006/jeth.1996.2218>
- CAMPBELL, J. Y. – LO, A. W. – MACKINLAY, C. A. [1997]: *The Econometrics of Financial Markets*. Princeton University Press. Princeton. <https://doi.org/10.1515/9781400830213-004>
- CHRISTOFFERSEN, P. F. [2012]: *Elements of Financial Risk Management*. Elsevier. New York.
- CLARK, P. C. [1973]: A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*. Vol. 41. No. 1. pp. 135–155. <http://dx.doi.org/10.2307/1913889>
- COLANDER, D. – KUPERS, R. [2014]: *Complexity and the Art of Public Policy – Solving Society’s Problems from the Bottom Up*. Princeton University Press. Princeton.
- EASLEY, D. – LÓPEZ DE PRADO, M. M. – O’HARA, M. [2012]: The volume clock: Insights into the high frequency paradigm. *The Journal of Portfolio Management*. Vol. 39. No. 1. pp. 19–29. <https://doi.org/10.3905/jpm.2012.39.1.019>
- HOLTON, G. A. [2003]: *Value at Risk: Theory and Practice*. Academic Press. Cambridge.
- JANSEN, S. [2020]: *Machine Learning for Algorithmic Trading: Predictive Models to Extract Signals from Market and Alternative Data for Systematic Trading Strategies with Python*. Packt Publishing. Birmingham.
- JONES, C. M. – KAUL, G. – LIPSON, M. L. [1994]: Transactions, volume and volatility. *Review of Financial Studies*. Vol. 7. Issue 4. pp. 631–651. <http://dx.doi.org/10.1093/rfs/7.4.631>
- JORION, P. [2006]: *Value at Risk: The New Benchmark for Managing Financial Risk*. Third Edition. McGraw-Hill. New York.
- KISS G. D. [2017]: *Volatilitás, extrém elmozdulások és tőkepiaci fertőzések*. Jatepress. Szeged.
- KOVÁCS L. – MARSÍ E. (szerk.) [2018]: *Bankmenedzsment – banküzemtan*. Magyar Bankszövetség. Budapest.
- LÓPEZ DE PRADO, M. M. [2018]: *Advances in Financial Machine Learning*. Wiley. New York.
- LEROY, S. F. [1989]: Efficient capital markets and martingales. *Journal of Economic Literature*. Vol. 26. No. 12. pp. 1583–1621.
- MANDELROT, B. [1963]: The variation of certain speculative prices. *The Journal of Business*. Vol. 36. No. 4. pp. 394–419. <http://dx.doi.org/10.1086/294632>
- MANDELROT, B. [1966]: Forecasts of future prices, unbiased markets, and “martingale” models. *The Journal of Business*. Vol. 39. No. 1. pp. 242–255. <http://dx.doi.org/10.1086/294850>
- MANDELROT, B. – TAYLOR, H. M. [1967]: On the distribution of stock price differences. *Operations Research*. Vol. 15. No. 6. pp. 1057–1062. <http://dx.doi.org/10.1287/opre.15.6.1057>
- MACKENZIE, D. [2006]: *An Engine, Not a Camera – How Financial Models Shape Markets*. MIT Press. Cambridge.

- MILLER, R. S. – SHORTER, G. [2016]: *High Frequency Trading: Overview of Recent Developments*. Congressional Research Service. <https://fas.org/sgp/crs/misc/R44443.pdf>
- SAMUELSON, P. [1965]: Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*. Vol. 6. No. 2. pp. 41–49.
- SAUNDERS, A. – CORNETT, M. [2018]: *Financial Markets and Institutions*. McGraw-Hill. New York.
- SAUNDERS, A. – CORNETT, M. [2021]: *Financial Institutions Management: A Risk Management Approach*. McGraw-Hill. New York.