



<input type="checkbox"/>	Kandidaatintutkielma
<input checked="" type="checkbox"/>	Pro gradu -tutkielma
<input type="checkbox"/>	Lisensiaatintutkielma
<input type="checkbox"/>	Väitöskirja

Oppiaine	Laskentatoimi ja rahoitus	Päivämäärä	8.7.2020
Tekijä	Juuso Marttinen	Sivumäärä	69 s. + liitteet
Otsikko	Osakeindeksituoton implisiittisen todennäköisyysjakauman estimointi – Normaalijakaumaoletuksen aiheuttamat hinnoitteluvirheet optiomarkkinoilla		
Ohjaaja	Prof. Luis H. R. Alvarez Esteban		

#### Tiivistelmä

Matemaattisen rahoituksen mallit ovat yksinkertaistuksia oikeiden markkinoiden toiminnasta. Stokastisten prosessien kohdalla tärkeimmät oletukset koskevat satunnaisuuden muotoa. Useimmiten satunnaisuutta kuvataan standardia normaalijakaumaa noudattavalla satunnaismuuttujalla sen laskennallisen helppouden vuoksi. Esimerkiksi optioiden hinnoittelussa yleisesti käytössä oleva Black–Scholes–Merton-malli pohjautuu geometriseen Brownin liikkeeseen, jossa kohde-etuuden hinnanmuutokset ovat log-normaalisti jakautuneita. Empiiriset tutkimukset ovat kuitenkin osoittaneet, että osakemarkkinoiden todellinen tuottojakauma ei ole normaali, minkä vuoksi Black–Scholes–Merton-mallissa esiintyy hinnoitteluvirheitä erityisesti toteutushinnan poiketessa reilusti nykyhinnasta.

Tämän tutkielman tavoitteena on estimoida markkinoilla vallitseva implisiittinen tiheysfunktio, jota optioiden tarjoama hintainformaatio heijastelee. Todennäköisyysjakauman muotoa ja sen eroavuutta normaalijakaumaan tarkastellaan usealla aikavälillä. Lopulta määritetään Black–Scholes–Merton-mallin aiheuttamat hinnoitteluvirheet eri optiokategorioissa.

Teoreettisessa viitekehyksessä käsitellään osakkeiden hintaprosesseja siinä tilanteessa, kun perinteisimmät oletukset ovat voimassa. Tarkastelu aloitetaan perinteisestä satunnaiskulkuhypoteesista, josta edetään stokastisiin differentiaaliyhtälöihin. Lyhyesti sivutaan myös pitkän muistin prosesseja. Osto- ja myyntioptioiden hinnoittelu perustetaan esiteltyihin kohde-etuuden hintaprosesseihin. Black–Scholes–Merton-mallin oletuksia ja niiden aiheuttamia virheitä lähestytään volatilitteettihymyn kautta. Lopulta esitellään myös periaate, jonka mukaan optioiden hinnat implikoivat markkinoilla vallitsevia tulevaisuudennäkymiä tiheysfunktion muodossa.

Empiirisessä osassa tutkitaan S&P 500 -indeksiä kohde-etuutenaan käyttäviä indeksioptioita ja estimoidaan niiden perusteella implisiittinen todennäköisyysjakauma. Menetelmänä käytetään ei-parametrista lähestymistapaa, jota tarkennetaan sovittamalla luonnollinen kuutiosplini volatilitteettihymyn. Tiheysfunktio estimoidaan optiohintafunktion toisesta derivaatasta toteutushinnan suhteen.

Tulokset osoittavat, että implisiittinen jakauma on sekä negatiivisesti vino että paksuhäntäinen. Eroavuus normaalijakaumaan on selvintä lyhyellä aikavälillä, kun taas pitkällä aikavälillä ero on vähäisempää ja paksuhäntäisyys näyttää jopa häviävän. Hinnoitteluvirhettä esiintyy myös eniten lyhyellä juoksuajalla. Pahimmillaan virhe on keskimäärin jopa satoja prosentteja, minkä vuoksi sijoittajat voivat altistua erittäin suurille virheille luottaessaan liikaa yhteen markkinoiden eniten käytetyistä hinnoittelumalleista.

Asiasanat	estimointi, hinnoittelu, indeksit, interpolointi, jakaumat, normaalijakauma, optiot, stokastiset prosessit, todennäköisyys
Muita tietoja	







**TURUN  
YLIOPISTO**  
Kauppakorkeakoulu

# OSAKEINDEKSITUOTON IMPLISIITTISEN TODENNÄKÖISYYSJAKAUMAN ESTIMOINTI

Normaalijakaumaoletuksen aiheuttamat hinnoitteluvirheet  
optiomarkkinoilla

Laskentatoimen ja rahoituksen  
pro gradu -tutkielma

Laatija:  
Juuso Marttinen

Ohjaaja:  
Prof. Luis H. R. Alvarez Esteban

8.7.2020  
Turku

Turun yliopiston laatujärjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -järjestelmällä.

# SISÄLLYS

<b>1 JOHDANTO</b> .....	7
1.1 Aiheen taustaa .....	7
1.2 Tutkielman tavoitteet ja analyysin lähtökohdat .....	8
<b>2 HINTARAKENTEET OSAKEMARKKINOILLA</b> .....	10
2.1 Satunnaisuus diskreetissä ajassa .....	10
2.2 Jatkuva-aikainen hintaprosessi .....	12
2.3 Hintaprosessin satunnaisuudessa esiintyvä aikariippuvuus .....	15
<b>3 OPTIOHINNOITTELU TUNNETULLA TODENNÄKÖI- SYYSJAKAUMALLA</b> .....	21
3.1 Option määritelmä ja hinnoittelu yksinkertaisilla markkinoilla .....	21
3.2 Yleisemmät diskreetit hinnoittelumallit .....	25
3.3 Optiohinnoittelu normaalijakaumaa käyttäen .....	28
3.4 Black–Scholes–Merton-mallin vääristymät oikeilla markkinoilla .....	31
3.5 Optioiden sisältämä informaatio tulevaisuudesta .....	34
<b>4 AINEISTON JA SOVELLETUN TUTKIMUSMENETELMÄN KUNNASTUS</b> .....	37
4.1 Aineiston hankinta ja alustavat muokkaustoimenpiteet .....	37
4.2 Menetelmän perusteet .....	38
4.3 Analyysin alkutoimenpiteet .....	39
4.4 Implisiittisen volatilitiitin määrittäminen ja tasoitus .....	41
4.5 Implisiittisen volatilitiitin hyödyntäminen hintojen tasoituksessa .....	46
4.6 Implisiittisen todennäköisyysjakauman estimointi .....	47
<b>5 TULOKSET JA JOHTOPÄÄTÖKSET</b> .....	48
5.1 Lyhyt aikaväli .....	52
5.2 Keskipitkä aikaväli .....	53
5.3 Pitkä aikaväli .....	54
5.4 Erot eri aikavälien kesken .....	56
5.5 Tulosten merkitys optioiden hinnoittelun kannalta .....	57

<b>6 YHTEENVETO</b> .....	62
<b>LÄHTEET</b> .....	65
<b>LIITTEET</b> .....	70
A    Analyysissä käytetyt ohjelmistot ja R-paketit .....	70
B    Epäonnistunut jakauman estimointi.....	71
C    Estimoitujen jakaumien momenttien ero normaalijakaumaan .....	72

## KUVIOT

Kuvio 1	Keskenään korreloituneita simuloituja Wiener-prosesseja .....	17
Kuvio 2	Fraktionaalisia Brownin liikkeitä eri Hurstin parametrin arvoilla	20
Kuvio 3	Yhden periodin binomipuu .....	22
Kuvio 4	Useamman periodin binomipuu .....	25
Kuvio 5	Volatiliteettihymyjä eri omaisuusluokissa .....	33
Kuvio 6	S&P 500 -indeksin annualisoidun osinkotuoton jakautuminen eri kuukausille.....	39
Kuvio 7	Luonnollinen kuutiosplini eri tasoitusparametrin arvoilla .....	43
Kuvio 8	Implisiittisen volatiliteetin tasoitus kuutiosplinillä .....	45
Kuvio 9	Implisiittisen volatiliteetin sovitteen mukaiset optiohinnat .....	46
Kuvio 10	Jakaumien eroavuuksien integroidut keskineliövirheet eri aikahorisonteilla .....	51
Kuvio 11	Lyhyen juoksuajan optioiden implikoima todennäköisyysjakauma	52
Kuvio 12	Keskipitkän juoksuajan optioiden implikoima todennäköisyysjakauma.....	54
Kuvio 13	Pitkän juoksuajan optioiden implikoima todennäköisyysjakauma	55
Kuvio 14	Prosentuaalinen hintavirhe optioiden hinnoissa.....	59

## TAULUKOT

Taulukko 1	Keskenään korreloituneiden Wiener-prosessien simulointi.....	16
Taulukko 2	Osto-optioiden palkkioita kohde-etuuden tilan mukaan .....	35
Taulukko 3	Tutkielmassa käytetyt aikasarja-aineistot .....	37
Taulukko 4	Estimoitujen jakaumien momenttien eroavuus normaalijakau- masta .....	50
Taulukko 5	Keskimääräinen integroitu keskineliövirhe eri aikahorisonteilla...	56
Taulukko 6	Optiohintojen prosentuaalinen keskivirhe eri kategorioissa .....	58



# 1 JOHDANTO

## 1.1 Aiheen taustaa

Matematiikka tarjoaa moderniin rahoitukseen lukuisia työkaluja, joilla markkinoiden toimintaa pystytään mallintamaan. Mallinnuksessa rakennetaan perusteltuihin oletuksiin pohjautuva kuva markkinoiden toiminnasta. Satunnaisissa eli stokastisissa prosesseissa tärkeimmät oletukset liittyvät satunnaisuuden käsittelyyn. Laskennallisen helppouden vuoksi satunnaisuutta kuvataankin yleensä standardilla normaalijakautuneella satunnaismuuttujalla, koska se on monesti myös teoreettisesti helposti perusteltavissa.

Mallien oletukset ovat yksinkertaistuksia oikeista markkinoista, minkä vuoksi niissä saattaa esiintyä virheitä. Tutkimuksen avulla voidaan aina kehitellä vaihtoehtoisia ratkaisuja, jotka korjaavat aiemman tiedon valossa esiintyneitä virheitä. Itse asiassa monet rahoituksen mallit toimisivat teorian muodostuksen kannalta satunnaisuuden muodosta riippumatta, mutta käytettävyyden nojalla usein turvaudutaan juuri normaalijakaumaan. Laskentatehon kasvaessa kynnys siirtyä monimutkaisempiin malleihin laskee jatkuvasti.

Normaalijakaumaan perustuva satunnaiskulkuhypoteesi alkoi muotoutua jo viime vuosisadan alussa, kun [Bachelier \(1900\)](#) tutki johdannaisten hintoja. Useita vuosikymmeniä myöhemmin [Godfrey, Granger ja Morgenstern \(1964\)](#) tekivät hypoteesista tunnetumman. Satunnaiskulussa prosessin muutokset ovat normaalisti jakautuneita ja toisistaan riippumattomia. Tämä on yhteydessä [Faman \(1970\)](#) tehokkaiden markkinoiden hypoteesin heikkojen ehtojen toteutumiseen, minkä mukaan tulevat hinnanmuutokset johtuvat vain uusista uutisista markkinoilla. Luvussa [2](#) nämä teorialat käsitellään tarkemmin ja niiden pohjalta edetään matemaattisen rahoituksen suosituimpien prosessien joukossa olevaan geometriseen Brownin liikkeeseen.

Yksi rahoituksen merkittävimmistä hinnoittelukaavoista syntyi 1970-luvun alkupuolella, kun [Black ja Scholes \(1973\)](#) sekä [Merton \(1973\)](#) johtivat eksplisiittisen kaavan optioiden hinnoitteluksi, mihin syvennytään tarkemmin luvussa [3](#). Mallin ansiosta optioiden kaupankäynti kiihtyi huomattavan paljon, koska markkinoiden osapuolilla oli käytettävissään työkalu, jonka lähes kaikki parametrit ovat yleisesti saatavilla markkinoilta ja siten objektiivisia kaikkien osapuolien välillä. Black–Scholes–Merton-kaavan taustalla oletetaan kohde-etuuden hinnan noudattavan geometrista Brownin liikettä, jossa satunnaisuus on normaalijakaumaan pohjautuvaa.

Empiirisissä tutkimuksissa on kuitenkin havaittu, että osakkeiden hinnanmuutokset eivät ole normaalisti jakautuneita (ks. esim. [Borowski 2018](#); [Liu, Moghaddam ja Serota 2019](#)). Tämä on yksi syy, miksi rahoituksessa varmasti yksi historian käyte-

tyimmistä malleista hinnoittelee optioita väärin. Luvussa 3.4 käsitellään syitä, miksi mallin oletukset poikkeavat oikeista markkinoista.

## 1.2 Tutkielman tavoitteet ja analyysin lähtökohdat

Entäpä jos markkinat olisivat aina oikeassa? Edellä viitattu Black–Scholes–Merton-malli kertoo kyllä option hinnan annetuilla parametreilla, mutta entä jos option hinta otetaankin markkinoilta ja ainut vapaa parametri, kohde-etuuden volatiliteetti, määritetään mallin perusteella. Tällöin puhutaan implisiittisestä volatiliteetista (engl. *implied volatility*). Normaalijakaumaoletuksen pitäessä paikkansa tulisi implisiittisen volatiliteetin olla vakio kaikilla toteutushinnoilla. Useimmiten implisiittinen volatiliteetti on kuitenkin epälineaarinen ja konvekssi funktio toteutushinnasta, jolloin puhutaan volatiliteettihymystä (engl. *volatility smile*).

Tämän tutkielman tavoitteena on haastaa osaketuottojen normaalijakaumaoletus ja selvittää, mitä informaatiota markkinoilla noteerattavat optioiden hinnat kertovat normaalijakauman pätevydestä. Tarkoitus on estimoida optiomarkkinoilla vallitseva käsitys tulevaisuudesta implisiittisen tiheysfunktion muodossa. Lisäksi tavoitteisiin lukeutuu implisiittisen jakauman muodon eroavuus normaalijakaumasta ja sen pysyvyys eri ajankohtina. Tutkielmassa pyritään myös selvittämään, onko estimoidussa implisiittisessä jakaumassa eroja eri aikaväleillä. Viimeisenä tavoitteena on nostaa esiin Black–Scholes–Merton-mallin tuottamat hinnoitteluvirheet, jotka johtuvat normaalijakauman hyödyntämisestä. Tiivistettynä tämän tutkielman tutkimuskysymykset ovat:

**Tutkimuskysymys 1:** Millainen on markkinoilla vallitseva implisiittinen todennäköisyysjakauma ja erityisesti poikkeako se normaalijakaumasta?

**Tutkimuskysymys 2:** Pysykö estimoidun implisiittisen todennäköisyysjakauman ja normaalijakauman ero stabiilina ajan suhteen?

**Tutkimuskysymys 3:** Millaisia eroja jakaumissa ilmenee ja mitkä tekijät niitä selittävät?

**Tutkimuskysymys 4:** Kuinka suuria hinnoitteluvirheitä syntyy, jos käytetään normaalijakaumaoletusta markkinoilla vallitsevan implisiittisen jakauman sijaan?

Luvussa 3.5 käsitellään, mihin tutkimusmenetelmä perustuu, ja puolestaan luvussa 4 esitellään tämän tutkielman menetelmälliset valinnat. Käytettävä metodologia perustuu Shimkon (1993) kehittämään menetelmään, mutta suositun tutkimusaiheen ansiosta tässä tutkielmassa hyödynnetään alan tutkimuksessa esiin nousseita tuoreempiakin näkökulmia. Tulokset esitetään luvussa 5.

Alkuperäinen tarkoitus oli tutkia Suomen osakemarkkinoita, mutta aineiston asettamien rajoitteiden vuoksi se osoittautui mahdottomaksi. Sen sijaan Yhdysval-

loista saatava aineisto on riittävän laadukasta ja kattavaa tutkielman toteuttamiseksi. Aineiston valinnassa on oltava tarkkana, koska esimerkiksi [Jackwerth ja Rubinstein \(1996, 1619\)](#) huomauttavat tutkimusmenetelmän olevan herkkä aineiston laajuudelle. Tästä syystä kattavampi aineisto on siis hyvin perusteltavissa.

Rahoituksen tutkimuksessa yleisestikin koetaan haastavaksi jotkin reaali maailman tekijät, jotka ovat läsnä markkinoilla. Myös tässä tutkielmassa joudutaan tekemään oletuksia markkinoiden toiminnasta joidenkin seikkojen suhteen. Ensinnäkin johdannaisten hinnoittelu vaatii arbitraasittomuutta, jotta ekvivalentti martingaalimittana on olemassa ja sitä kautta voidaan hyödyntää riskineutraalia hinnoittelua. Tähän oletukseen palataan vielä tarkemmin luvussa 3. Toiseksi markkinoiden tehokkuudesta on oletettava, että kaupankäyntikustannuksia tai veroja ei ole, koska niiden huomioon ottaminen ei olisi tämän tutkielman kannalta merkittävää tai monimutkaisuuden vuoksi myöskään kannattavaa. Kolmanneksi jotta voitaisiin kuvata jatkuvan jakauman tiheysfunktioita, tarvittaisiin tila-avaruutta kuvaamaan toteutushintojen jatkumo. Tällaista aineistoa ei luonnollisestikaan ole saatavilla, joten jakauma on estimoitava diskreetisti, vaikkakin estimoinnin tarkkuuteen voidaan vaikuttaa jakovälin valinnalla.

Tarkasteluaikaväliksi on valittu mahdollisimman tuore aineisto työn kirjoitusai-kataulu huomioiden. Aineisto alkaa vuoden 2018 loppupuolelta ja päättyy vuoden 2020 maaliskuuhun. Tämä mahdollistaa useamman aikahorisontin hyödyntämisen ja mahdollisten erojen löytämisen.

Tämän tutkielman pääpaino on erilaisten matemaattisten ja tilastotieteellisten työkalujen soveltamisessa eikä niinkään teoreettisessa käsittelyssä. Stokastiikan ja todennäköisyysteorian tuloksia ei siksi johdeta, ja matemaattisten muodollisuuksien tarkkuudesta tingitään paikoittain, koska tarkoituksena on hyödyntää valmiita työkaluja soveltuvien osien eikä johtaa uusia matemaattisia tuloksia.

## 2 HINTARAKENTEET OSAKEMARKKINOILLA

### 2.1 Satunnaisuus diskreetissä ajassa

Matemaattisen rahoituksen perusteet alkoivat muotoutua, kun [Bachelier \(1900\)](#) tutki johdannaisten hintojen kehitystä. Hänen väitöskirjassaan tutkitaan odotusarvon käsitettä ja todennäköisyysrakenteita kahdesta näkökulmasta. Ensiksi käsitellään olemassa olevan tiedon nojalla muodostettua odotusarvoa tulevasta kassavirrasta. Toisin sanoen markkinoiden oletetaan siis noudattavan determinististä hintadynamiikkaa. Markkinoiden tehokkuuden näkökulmasta [Bachelier \(1900\)](#) määrittelee, että markkinoilla todennäköisimpänä pidetyn hinnan tulee olla juuri deterministisesti laskettavissa oleva nykyarvo. Muussa tapauksessa markkinoilla esiintyisi hinnoitteluvirhe, joka korjaantuisi itsestään.

Toiseksi [Bachelier \(1900\)](#) tutkii tulevaisuuden satunnaisten tapahtumien aiheuttamia muutoksia aiemmin deterministisenä pidettyyn hintaprosessiin. Tämän tulevaisuudessa sijaitsevan odotusarvon määrittystä [Bachelier](#) pitää erityisen merkityksellisenä markkinoilla toimivien osapuolien näkökulmasta ja osoittaa todennäköisyyslaskennan teorian turvin, että myös satunnaisuuden vallitessa odotusarvo on sama kuin nykyarvo. Toisin sanoen tulevaisuuden satunnaisten hinnanmuutosten odotusarvo on nolla. Lisäksi [Bachelier \(1900\)](#) osoittaa jakauman olevan gaussinen ja sen keskihajonnan olevan ajan neliöjuuren funktio. Tästä syystä jakauma on siis symmetrinen, minkä takia hinnan nousun ja laskun todennäköisyydet ovat samat.

Myöhemmin muun muassa [Mandelbrot \(1963, 394–395\)](#) havaitsi, ettei 1900-luvun alkupuolen empiiriset havainnot tue hintojen normaalijakautuneisuutta, vaan empiirinen jakauma on huipukkaampi ja hänniltään pidempi. Hänen mukaansa syy eroavuuteen on toisen momentin epänormaalit vaihtelut. [Bachelierin](#) väitöskirjaa väheksymättä [Mandelbrot \(1963\)](#) esittää, että normaalijakautuneisuuden tulisi olla voimassa logaritmistien hintojen differensseille eli jatkuva-aikaisille tuotoille, vaikkakin hänen mukaansa normaalijakaumaa parempi jakauma olisi yleisempi vakaa Pareto-jakauma. Kuitenkin logaritmuunnos estää esimerkiksi jakauman vasemman hännän ulottumisen negatiiviseksi eli osakemarkkinoille sovellettuna hinnat eivät voi jakauman mukaan olla negatiivisia. Klassisen todennäköisyysteorian näkökulmasta katsottuna osakemarkkinoiden log-normaalijakautuneisuus on siis matemaattisesti hyvin perusteltu oletus.

Muiden muassa [Godfrey ym. \(1964\)](#) toivat tunnetuksi satunnaiskulkuna (engl. *random walk*) tunnettua hintaprosessia. Itse prosessi voidaan määrittellä matemaattisesti niin sanottuna ARIMA(0, 1, 0)-prosessina hinnan logaritmille

$$r_t := \Delta(\ln S_t) = \ln S_t - \ln S_{t-1} = \ln \left( \frac{S_t}{S_{t-1}} \right) = \sigma Z_t, \quad (2.1)$$

missä  $r_t$  on jatkuva-aikainen kohde-etuuden tuotto,  $S_t$  on kohde-etuuden hinta hetkellä  $t$ ,  $\sigma$  on skaalaustekijä (tuoton volatilitiitti) ja  $Z_t \sim \text{NID}(0, 1)$  on satunnaisuutta ajava stokastinen prosessi, joka noudattaa standardia normaalijakaumaa ja jolla on toisistaan riippumattomat realisaatiot. Osakemarkkinat ovat kuitenkin pitkällä aikavälillä osoittaneet olevansa jatkuvassa nousussa (ks. esim. [Investopedia 2020](#)), joten edellä kuvattuun satunnaiskulkuprosessiin (2.1) on syytä lisätä deterministinen vakio, joka kuvaa markkinan keskimääräistä kasvuvauhtia. Satunnaiskulkuprosessi ajautumatermillä (engl. *drift*) on muotoa

$$r_t = \mu + \sigma Z_t. \quad (2.2)$$

Sen muodostaman logaritmisen hintaprosessin odotusarvo välillä  $[0, t]$  ja alkuhinnalla  $S_0$  on

$$\mathbb{E}[\ln S_t] = \mathbb{E}[\ln S_0 + r_1 + r_2 + \cdots + r_t] = \ln S_0 + \mu t + \sigma \mathbb{E} \left[ \sum_{j=1}^t Z_j \right].$$

Koska kaikki  $Z_j$ :t ovat toisistaan riippumattomia ja odotusarvo on lineaarinen operaatio, voidaan odotusarvo-operaattori viedä summan sisäpuolelle, jolloin tuoton odotusarvoksi tulee

$$\mathbb{E}[r_t] = \mathbb{E}[\ln S_t] - \ln S_0 = \mu t + \sigma \sum_{j=1}^t \mathbb{E}[Z_j] = \mu t + \sigma \sum_{j=1}^t 0 = \mu t$$

Satunnaiskulkuprosessin varianssi voidaan myös johtaa suoraan prosessiyhtälöstä:

$$\text{Var}[\ln S_t] = \text{Var}[\ln S_0 + r_1 + r_2 + \cdots + r_t] = \text{Var} \left[ \ln S_0 + \mu t + \sigma \sum_{j=1}^t Z_j \right].$$

Determinististen vakioiden ( $S_0$ ,  $\mu$  ja  $t$ ) varianssin ollessa 0 ja vedoten samaan riippumattomuusoletukseen kuin edellä odotusarvon yhteydessä saadaan prosessin varianssi muotoon

$$\text{Var}[\ln S_t] = \sigma^2 \sum_{j=1}^t \text{Var}[Z_j] = \sigma^2 \sum_{j=1}^t 1 = \sigma^2 t$$

eli satunnaiskulkuprosessin varianssi kasvaa ajan funktiona räjähtäen lopulta äärettömäksi, kun  $t \nearrow \infty$ . Sekä odotusarvon että varianssin aikariippuvuudesta johtuen ajautuva satunnaiskulkuprosessi ei selvästikään ole heikosti stationaarinen prosessi.

[Cox, Ross ja Rubinstein \(1979\)](#) olivat ensimmäisiä, jotka tutkivat optioiden hinnoittelua binomipuun avulla. Binomipuussa jokaisella periodilla on kaksi mahdollista maailmantilaa  $(\omega_1, \omega_2)$ , joihin sidotut realisaatiot ovat tunnettuja. Lyhennettäessä periodien pituutta periodien määrä kasvaa. Keskeiseen raja-arvolauseeseen nojaten riippumattomien havaintojen summan jakauma lähestyy normaalijakaumaa

riippumatta yksittäisen havainnon jakaumasta, kun havaintojen määrä lähestyy ääretöntä (Ross 2011, 30). Havaittu tulos pätee mille tahansa instrumentille, jonka hintaprosessia voidaan mallintaa binomipuulla (Cox ym. 1979) tai yleisemmin millä tahansa muulla jakaumalla, kunhan eri periodien realisaatiot generoituvat toisistaan riippumattomasti samasta jakaumasta sekä keskiarvo ja varianssi ovat äärellisenä olemassa. Myös tästä näkökulmasta osaketuottojen pitkän aikavälin normaalijakautuneisuus on perusteltu oletus. Luvussa 3.1 palataan vielä tarkemmin optioiden hinnoitteluun binomiprosessin avulla.

Fama (1970) tutki informaation vaikutusta hintojen kehitykseen ja kehitti yhden kaikkien aikojen tunnetuimmista rahoitusteorioista: tehokkaiden markkinoiden hypoteesin (engl. *efficient market hypothesis*). Heikon tehokkuuden ehdot täyttävillä markkinoilla osakkeen historiallisesta hintakehityksestä ei voida päätellä tulevaa hintaa. Keskivahvan tehokkuuden ehdot täyttävillä markkinoilla osakkeen hinta sisältää kaiken julkisesti saatavilla olevan informaation. Vahvan tehokkuuden ehdot täyttävillä markkinoilla kaikki olemassa oleva informaatio on hinnoiteltu osakkeen hintaan. Implisiittisesti on myös olemassa tilanne, jossa markkinat eivät täytä edes heikon tehokkuuden ehtoja. Tällöin historiallinenkaan tieto ei ole täysimääräisenä mukana osakkeen hinnassa.

Markkinoiden tehokkuus vaikuttaa oleellisesti satunnaiskulkuprosessin sovellettavuuteen empiirisellä aineistolla. Tavanomainen odotusarvoon perustuva seuraavan periodin hinta on Faman (1970) mukaan

$$\mathbb{E}[S_{t+1}|\Omega_t] = (1 + \mathbb{E}[r_{t+1}|\Omega_t])S_t, \quad (2.3)$$

missä  $\Omega_t$  on hetkellä  $t$  olemassa oleva informaatio. Tehokkailla markkinoilla informaatio  $\Omega_t$  on kaikilla hetkillä  $t$  mukana hinnassa, joten tuoton odotusarvo on aina ajasta riippumaton, koska hinnan muutos riippuu vain uudesta informaatiosta hetkellä  $t + 1$ . Tehokkailla markkinoilla hintojen muutokset ovat siis aina täysin satunnaisia ja satunnaiskulkuprosessin (2.1) mukaan myös riippumattomia toisistaan. (Fama 1970.)

## 2.2 Jatkuva-aikainen hintaprosessi

Osakkeiden hintojen voidaan perustellusti olettaa noudattavan Markov-prosessia. Tässä kontekstissa Markov-ominaisuus tarkoittaa sitä, että osakkeen tuleva hinta riippuu ainoastaan sen nykyisestä hinnasta. Mennyt hintahistoria ei siis vaikuta hinnan tulevaan kehityspolkuun, mikä vastaa sitä, että Faman (1970) määrittelemät heikon tehokkuuden ehdot ovat voimassa osakemarkkinoilla. (Hull 2011, 280–281.) Yksi tehokkaiden markkinoiden hypoteesin puolustajista on Malkiel (2019, 2003), vaikka myös erilaisille osakemarkkinoiden anomalioiden kuten tammikuuilmiölle ja

momentum-strategialle, jotka vastustavat heikkojen ehtojen toteutumista, on ajan saatossa löydetty tukea (ks. esim. Keim 1985; Jegadeesh ja Titman 1993).

Ajautuvassa satunnaiskulkuprosessissa (2.2) oleva satunnaisuutta ajava gaussinen valkoisen kohinan prosessi  $\{Z_t\}$  muodostaa Wiener-prosessin sillä ehdolla, että prosessi alkaa nollassa. Valkoisen kohinan ominaisuudet takaavat suoraan Wiener-prosessin vaatimat riippumattomat lisäykset ja jatkuvuuden. Jos ajanhetkille pätee  $r < s \leq t < u$ , Wiener-prosessin määritelmä on

$$\begin{aligned} W_0 &= 0 \\ (W_u - W_t) &\perp (W_s - W_r) \\ (W_t - W_s) &\sim N(0, t - s) \end{aligned} \tag{2.4}$$

eli sillä on satunnaiskulustakin tuttu ominaisuus, että prosessin varianssi on yhtä suuri kuin aikavälin pituus. (Ross 2011; Björk 2009.)

Soveltamalla ajautuvaa satunnaiskulkuprosessia (2.2) tuoton sijaan suoraan osakkeen hinnan differenssiin ja käyttämällä Wiener-prosessia (2.4) osakkeen hintaprosessia voidaan kuvata differenssiyhtälöllä

$$\Delta S_t = \mu(t, S_t) \Delta t + \sigma(t, S_t) \Delta W_t, \tag{2.5}$$

missä aikavälin  $[t, t + \Delta t]$  pituus on  $\Delta t$  sekä  $\mu(\cdot)$  ja  $\sigma(\cdot)$  ovat ajanhetkestä  $t$  ja osakkeen hinnasta  $S_t$  riippuvat kasvu- ja diffuusiokertoimet. Tavanomaisen differentiaalilaskennan mukaan olisi houkuttelevaa jakaa yhtälö (2.5) puolittain välin pituudella  $\Delta t$  ja antaa aikavälin lähestyä nollaa  $\Delta t \searrow 0$ . Tämä ei kuitenkaan ole mahdollista, koska vaikka Wiener-prosessi on jatkuva, se ei lähes varmasti ole missään differentioituva. (Björk 2009, 41.) Tarvitaan siis muita keinoja stokastisten differentiaaliyhtälöiden ja integraalien käsittelyyn.

Wiener-prosessin ei-differentioituvuuden takia integrointia on lähestyttävä toisesta näkökulmasta. Antamalla differenssiyhtälön (2.5) aikavälin lähestyä nollaa  $\Delta t \searrow 0$  ilman, että sitä ensin jaetaan välin pituudella  $\Delta t$ , osakkeen hintaprosessin stokastiseksi differentiaaliyhtälöksi saadaan

$$dS_t = \mu(t, S_t) dt + \sigma(t, S_t) dW_t, \tag{2.6}$$

jonka ensimmäinen termi kuvaa determinististä keskimääräistä tuottoa ja jälkimmäinen termi satunnaisuuden aiheuttamaa volatilitteettia. (Björk 2009, 42.) Yhtälö (2.6) noudattaa geometrista Brownin liikettä silloin, kun kasvu- ja diffuusiokertoimet ovat suhteellisia osuuksia osakkeen hinnasta eli  $\mu(t, S_t) = \mu_t S_t$  ja  $\sigma(t, S_t) = \sigma_t S_t$ . Tässä muodossa yhtälö (2.6) on matemaattisen rahoituksen yksi merkittävimmistä rakennuspalikoista.

Stokastinen differentiaaliyhtälö (2.6) voidaan esittää integraalimuodossa seuraavasti:

$$S_t = S_0 + \int_0^t \mu(t, S_t) dt + \int_0^t \sigma(t, S_t) dW_t, \tag{2.7}$$

missä  $S_0$  kuvaa osakkeen hintaa alkuhetkellä. Ensimmäinen integraali  $dt$ :n suhteen on luonteeltaan kokonaan deterministinen, joten sille pätee perinteisen Riemann-integraalin säännöt. Sen sijaan jälkimmäinen integraali  $dW_t$ :n suhteen on stokastinen, minkä takia sen käsittelyyn tarvitaan Itô'n (1944) integraalia. (Björk 2009, 42.)

Itô–Döblinin lauseena tunnettu tulos on matemaattisen rahoituksen varmasti yksi kuuluisimmista ja käytetyimmistä tuloksista. Döblin (ks. Bru ja Yor 2002) ja Itô (1946) osoittavat, että stokastisella integraalilla (2.7) on olemassa vain yksi ratkaisu. Soveltamalla lausetta kahdesti jatkuvasti differentioituvaan kuvaukseen voidaan prosessin  $f(t, S_t)$  dynamiikkaa kuvaava stokastinen differentiaaliyhtälö esittää muodossa

$$df(t, S_t) = \left( \frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial S} + \frac{1}{2} \sigma^2 \frac{\partial^2 f}{\partial S^2} \right) dt + \sigma \frac{\partial f}{\partial S} dW_t. \quad (2.8)$$

Koska tässä tutkielmassa tarkastelun kohteena ovat jatkuva-aikaiset tuotot, kuten yhtälön (2.2) yhteydessä todettiin, on luontevaa soveltaa Itô–Döblinin lausetta kuvaukseen  $f(S_t) = \ln S_t$ . Tällöin lauseen (2.8) soveltamisessa tarvittavat osittaisderivaatat ovat

$$\begin{aligned} \frac{\partial}{\partial t} \ln x &= 0 \\ \frac{\partial}{\partial x} \ln x &= \frac{1}{x} \\ \frac{\partial^2}{\partial x^2} \ln x &= -\frac{1}{x^2} \end{aligned}$$

ja näin ollen sovellus kokonaisuudessaan johtaa tulokseen

$$\begin{aligned} d(\ln S_t) &= \left( 0 + \mu S_t \frac{1}{S_t} + \frac{1}{2} \sigma^2 S_t^2 \frac{-1}{S_t^2} \right) dt + \sigma S_t \frac{1}{S_t} dW_t \\ &= \left( \mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dW_t. \end{aligned} \quad (2.9)$$

Integroimalla (2.9) nyt välin  $[t, T]$  yli saadaan alkuperäinen hintaprosessin differentiaaliyhtälö (2.6) ratkaistua yleiseen muotoon

$$\begin{aligned} \ln S_T - \ln S_t &= \left( \mu - \frac{1}{2} \sigma^2 \right) (T - t) + \sigma (W_T - W_t) \\ \Leftrightarrow S_T &= S_t e^{(\mu - \frac{1}{2} \sigma^2)(T-t) + \sigma(W_T - W_t)} \end{aligned} \quad (2.10)$$

Seuraavaksi huomataan Wiener-prosessin (2.4) ominaisuuksista, että satunnaisuutta ajava Wiener-prosessin muutos on normaalisti jakautunut  $(W_T - W_t) \sim N(0, T - t)$ . Tämän ansiosta yhtälöä (2.10) on varsin helppo käsitellä normaali-jakauman momentit generoivan funktion avulla, koska Wiener-prosessin muutos  $W_T - W_t$  on samoin jakautunut kuin tavanomainen normaalisti jakautunut satunnaismuuttuja, jonka varianssi on aikavälin pituus  $T - t$ .

### 2.3 Hintaprosessin satunnaisuudessa esiintyvä aikariippuvuus

Vaikka teoreettisesti edellä osoitettiin normaalijakauman olevan validi oletus osakemarkkinoiden tuoton jakaumasta, ei osakemarkkinoiden tuotto empiirisesti tarkasteltuna kuitenkaan ole yleensä normaalisti jakautunut (ks. esim. [Borowski 2018](#); [Liu ym. 2019](#)). Markkinoiden mallinnus normaalijakaumaa hyödyntäen ei onnistu kovin tarkasti jakauman suhteen kriittisten mallien osalta, kuten esimerkiksi optioiden hinnoittelussa käytetyllä Black–Scholes–Merton-kaavalla ([Yan ja Han 2019](#)), jota käsitellään myöhemmin tämän tutkielman luvussa 3.3.

Empiirisen jakauman poikkeaminen normaalijakaumasta voi tietysti johtua monestakin syystä, mutta yksi merkittävä tekijä on kaavassa (2.4) esille nostettu Wiener-prosessin riippumattomuus ajasta. Oletus riippumattomuudesta on vahva, mikä toisaalta helpottaa mallinnusta, mutta myös tulee helposti rikotaksi, kun tarkastellaan osakemarkkinoita empiirisesti.

Korreloitunut Wiener-prosessi muodostuu tavanomaisista riippumattomista Wiener-prosesseista  $W_1, \dots, W_d$  ja kerroinmatriisista

$$\boldsymbol{\delta} = \begin{bmatrix} \delta_{11} & \delta_{12} & \cdots & \delta_{1d} \\ \delta_{21} & \delta_{22} & \cdots & \delta_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{n1} & \delta_{n2} & \cdots & \delta_{nd} \end{bmatrix}$$

niin, että jokaisen rivivektorin pituus eli euklidinen normi  $\|\delta_i\| = 1, \forall i$ . Matriisitulona  $n$ -ulotteinen korreloitunut Wiener-prosessi on muotoa

$$\bar{\mathbf{W}} = \boldsymbol{\delta} \mathbf{W}$$

ja alkiokohtaisesti siis

$$\bar{W}_i = \sum_{j=1}^d \delta_{ij} W_j, \quad \forall i$$

joista jokainen on erillään tavanomainen riippumaton Wiener-prosessi, vaikka prosessit ovat keskenään korreloituneita. Näin muodostettujen Wiener-prosessien korrelaatiomatriisiksi tulee

$$\boldsymbol{\rho} = \boldsymbol{\delta} \boldsymbol{\delta}^T.$$

([Björk 2009](#), 59–60.) Korreloituneet Wiener-prosessit ovat yksinkertaistaen siis lineaarikombinaatioita tavanomaisista riippumattomista Wiener-prosesseista. Näin olleiden korreloituneiden Wiener-prosessien taustalla olevat valkoisen kohinan prosessit ovat edelleen IID-prosesseja, koska normaalisti jakautuneiden satunnaismuuttujien summa on edelleen normaalisti jakautunut.

Björkiä (2009, 61–62) mukaillen havainnollistetaan korreloituneita Wiener-prosesseja yksinkertaisen esimerkin kautta: käytetään kahta tavanomaista eli riippumatonta Wiener-prosessia  $W^d$ ,  $d = 1, 2$  ja käytetään kerroinmatriisin  $\boldsymbol{\delta}$  tilalla helpommin omaksuttavampaa kovarianssimatriisia muistuttavaa yhteisvaihtelumatriisia

$$\boldsymbol{\sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \vdots & \vdots \\ \sigma_{n1} & \sigma_{n2} \end{bmatrix}$$

jolloin korreloituneet Wiener-prosessit ovat muotoa

$$\bar{W}^i = \frac{1}{\|\boldsymbol{\sigma}_i\|} \boldsymbol{\sigma}_i \mathbf{W}, \quad \forall i$$

missä  $\boldsymbol{\sigma}_i$  on  $i$ :s rivi matriisista  $\boldsymbol{\sigma}$ . Näitä korreloituneita Wiener-prosesseja sovelletaan nyt käyttämällä pohjana stokastista differentiaaliyhtälöä (2.6)

$$dS_t^i = \mu^i S_t^i dt + \|\boldsymbol{\sigma}_i\| S_t^i d\bar{W}^i, \quad \forall i$$

Kuviossa 1 on piirrettynä kaksi riippumatonta Wiener-prosessia (mustat käyrät) ja niiden perusteella viisi erilaista korreloitunutta Wiener-prosessia. Simulaatiossa käytetyt parametrit on esitetty taulukon 1 osassa (a). Osat (b) ja (c) kuvaavat prosessien laskennallista ja erään simulaation kohdalla toteutunutta korrelaatiomatriisia.

**Taulukko 1:** Keskenään korreloituneiden Wiener-prosessien simulointi

(a) Simuloinnissa käytetyt parametrit

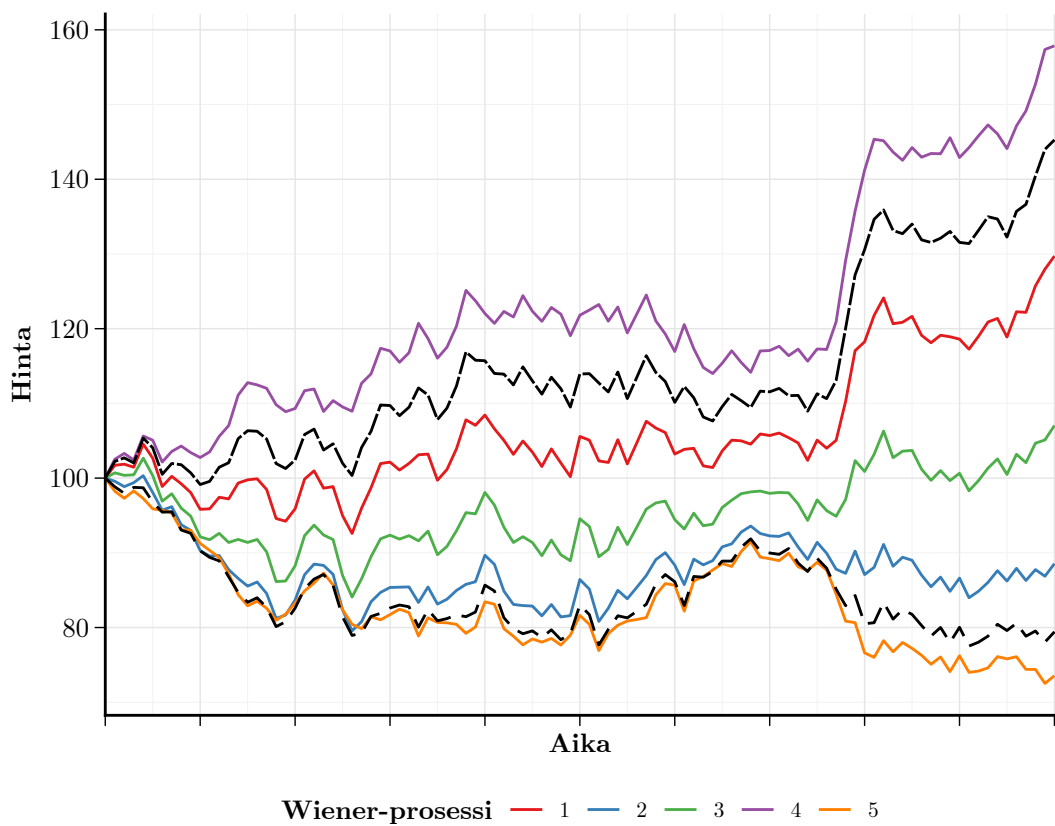
Muuttuja	Käytetty arvo
$\mu$	10 %
$\sigma$	20 %
$T$	1
$dt$	$\frac{1}{100}$
$\boldsymbol{\sigma}$	$\begin{bmatrix} 15 \% & 5 \% \\ 5 \% & 15 \% \\ 10 \% & 10 \% \\ 30 \% & -10 \% \\ -10 \% & 30 \% \end{bmatrix}$

(b) Laskennallinen korrelaatiomatriisi

$\rho$	1	2	3	4	5
1	1	0,600	0,894	0,800	0,000
2	0,600	1	0,894	0,000	0,800
3	0,894	0,894	1	0,447	0,447
4	0,800	0,000	0,447	1	-0,600
5	0,000	0,800	0,447	-0,600	1

(c) Erään simulaation toteutunut korrelaatiomatriisi

$\tilde{\rho}$	1	2	3	4	5
1	1	0,684	0,914	0,782	0,147
2	0,684	1	0,921	0,079	0,822
3	0,914	0,921	1	0,461	0,536
4	0,782	0,079	0,461	1	-0,502
5	0,147	0,822	0,536	-0,502	1



**Kuvio 1:** Keskenään korreloituneita simuloituja Wiener-prosesseja

Erityisesti kuvion 1 loppupäässä, kun generoinnin pohjalla käytetyt riippumattomat prosessit lähtevät erkanemaan, havaitaan hyvin, miten uudet prosessit ovat korreloituneet toistensa kanssa. Esimerkiksi prosessit  $\bar{W}^3$  ja  $\bar{W}^4$  käyttäytyvät hyvin samalla tavalla prosessin  $\bar{W}^1$  kanssa, minkä takia niiden väliset toteutuneet korrelaatiot ovat hyvin korkeita ( $\tilde{\rho}_{13} = 0,914$  ja  $\tilde{\rho}_{14} = 0,782$ ). Samoin prosessi  $\bar{W}^2$  ja  $\bar{W}^3$  ovat hyvin vahvasti korreloituneet ( $\tilde{\rho}_{23} = 0,921$ ).

Keskenään korreloituneet Wiener-prosessit toimivat tässä tutkielmassa lähinnä pohjustavana konseptina, koska tarkastelun kohteena on yksi kohde-etuus eikä useampia eri kohde-etuuksia, jotka olisivat riippuvia toisistaan. Keskinäinen korrelaatio tuo kuitenkin esiin sen tärkeän näkökulman, että Wiener-prosessista puhuttaessa yleensä aina keskitytään sen IID-ominaisuuteen, vaikka se voi yhtä hyvin sisältää riippuvuuksia sekä toisista Wiener-prosesseista että omasta historiastaan. Tämän tutkielman kannalta merkittävämpää on keskittyä yksiulotteiseen satunnaisprosessiin, jossa riippuvuus voi ilmentyä aikaulottuvuudessa eli tavanomaisen Wiener-prosessin generoiva gaussinen valkoinen kohina ei tällöin olekaan enää IID-prosessi. Kohina on tällaisessa tapauksessa autokorreloitunutta. Tällaiseen malliluokkaan siirryttäessä puhutaan fraktionaalisesta prosessista (engl. *fractional process*).

Mandelbrot ja Van Ness (1968, 423–424) toivat ensimmäisinä tunnetuksi fraktionaalisen Brownin liikkeen (engl. *fractional Brownian motion*, fBm) rahoitussovelluksiin. Kyse on Wiener-prosessista, missä parametri  $H \in (0, 1)$ , jota kutsutaan myös Hurstin parametriksi, määrittää integroituneisuuden tason. Fraktionaalinen Brownin liike on määritelmällisesti

$$W_t^H - W_0^H = \frac{1}{\Gamma(H + \frac{1}{2})} \left\{ \int_{-\infty}^0 [(t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}}] dW_s + \int_0^t (t-s)^{H-\frac{1}{2}} dW_s \right\} \quad (2.11)$$

missä  $\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$  on gammafunktio. Prosessi voidaan jakaa parametrin  $H$  suhteen kolmeen eri osaan: (1)  $H = \frac{1}{2}$ , jolloin prosessi vastaa tavanomaista Wiener-prosessia, (2)  $H \in (0, \frac{1}{2})$ , jolloin prosessin lisäyksillä on negatiivinen autokorrelaatio ja (3)  $H \in (\frac{1}{2}, 1)$ , jolloin prosessin lisäyksillä on positiivinen autokorrelaatio.

Bender, Sottinen ja Valkeila (2011) käsittelevät fraktionaalista Brownin liikettä ja listaavat muutamia prosessin tyypillisiä ominaisuuksia. Sen kovarianssirakenne on muotoa

$$\text{Cov}[W_t, W_s] = \frac{1}{2} [t^{2H} + s^{2H} - (t-s)^{2H}], \quad t \geq s.$$

Vaikka fraktionaalinen prosessi itse ei ole stationaarinen, sen lisäykset puolestaan ovat stationaarisia aivan kuten tavanomaisessakin Wiener-prosessissa. Lisäksi vii-

meisessä edellä mainituista kategorioista eli kun  $H \in (\frac{1}{2}, 1)$  fraktionaalisella Brownin liikkeellä on pitkä muisti

$$\rho_n = \text{Cov}[W_t^H - W_{t-1}^H, W_{t+n}^H - W_{t+n-1}^H], \quad \sum_{n=1}^{\infty} |\rho_n| = \infty,$$

kun tavanomaisella Wiener-prosessilla lisäykset ovat toisistaan täysin riippumattomat. Hurstin parametrin  $H$  vaikutusta muistiin on simuloitu R:llä `somebm`-paketin `fbm`-funktioilla aikavälillä  $t \in [0, 1]$  käyttämällä 200 osaväliä. Demonstraatio on kuvattu kuviossa 2, jonka osassa (a) on piirretty prosessien polut ja osassa (b) niiden autokorrelaatiofunktiot.

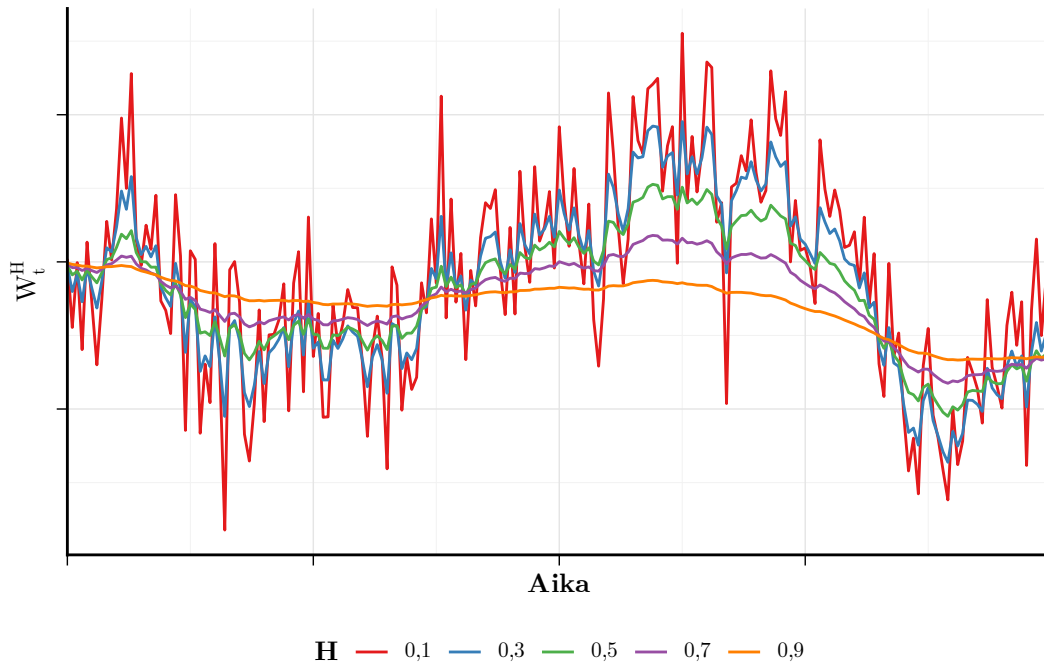
Vihreä käyrä ( $H = 0,5$ ) vastaa tavanomaista geometrista Brownin liikettä. Sen autokorrelaatiofunktio osoittaa, että prosessin lisäykset eivät ole tilastollisesti merkitseviä millään viiveellä 30:een saakka<sup>1</sup>. Negatiivista autokorrelaatiota esiintyy pienillä  $H$ :n arvoilla, mutta tässä simulaatiossa kuitenkin vain ensimmäisellä viiveellä. Negatiiviselle autokorrelaatiolle tyypillisesti polut sahaavat huomattavasti normaalia Wiener-prosessia enemmän.

Suurimpana kiinnostuksen kohteena on kuitenkin prosessit, joiden  $H > 0,5$ , koska niillä on pitkän muistin ominaisuuksia. Kun  $H = 0,7$ , ei autokorrelaatio vielä ole erityisen vahvaa, mutta selvästi autokorrelaatiokertoimet vaimenevat hitaasti<sup>2</sup>. Selkein tapaus on  $H = 0,9$ , jossa autokorrelaatiofunktio vaimenee äärimmäisen hitaasti ja osoittaa, että satunnaisuudella on erittäin vahva riippuvuus omasta historiastaan. Ero on radikaali verrattuna tavalliseen Wiener-prosessiin. Tämä voidaan havaita myös käyrien muodoista, koska prosessien polut ovat huomattavasti sileämpiä. Näin ollen tavallisimmin sovelletun differentiaaliyhtälön (2.9) taustalla oleva riippumattomuusoletus rikkoutuu selvästi.

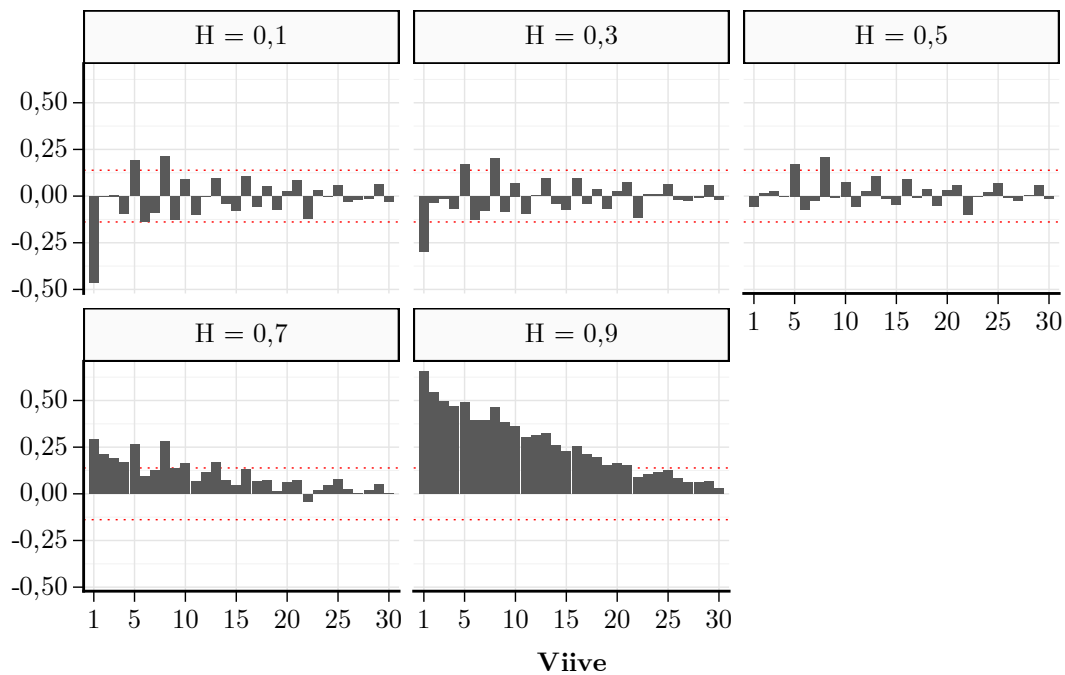
Aiemmassa kirjallisuudessa muun muassa [Jamdee ja Los \(2007\)](#) nostavat esiin pitkän muistin vaikutuksen optioiden arvostukseen. Lisäksi heidän kirjallisuuskatsauksensa esittelee lukuisia tutkimuksia, joissa Hurstin parametrin on todettu olevan yli 0,5. Myös esimerkiksi [Gomes, Soares, Gama ja Matos \(2018\)](#) löytävät merkkejä pitkästä muistista useilla eurooppalaisilla osakeindekseillä Hurstin parametrin arvojen asettuessa välille 0,55–0,6 selitysasteiden ollessa erittäin korkeita. Tulokset ovat linjassa myös heidän käsittelemiensä aiempien tutkimusten kanssa. He huomauttavat kuitenkin, että pitkän muistin vaikutus on vähentynyt ajan kuluessa.

<sup>1</sup> Satunnaisuudesta ja pienehköstä havaintomäärästä johtuen joidenkin yksittäisten viiveiden autokorrelaatiokertoimet ylittävät approksimatiiviset kriittiset rajat. Ylityksiä ei tapahdu, jos otoskoko kasvatetaan ( $N = 2000$ ).

<sup>2</sup> Käyttämällä tiheämpää osavälijakoa ( $N = 2000$ ) vaimeneminen tulee vielä selvemmin ilmi.



(a) Prosessien polkurealisaatiot

(b) Prosessien lisäysten autokorrelaatiofunktiot eri  $H$ :n arvoilla

Huomaa, että tilastollisen merkitsevyyden kriittiset rajat punaisella katkoviivalla ovat approksimatiivisia.

**Kuvio 2:** Fraktionaalisia Brownin liikkeitä eri Hurstin parametrin arvoilla

### 3 OPTIOHINNOITTELU TUNNETULLA TODENNÄKÖISYYSJAKAUMALLA

#### 3.1 Option määritelmä ja hinnoittelu yksinkertaisilla markkinoilla

Optiot ovat johdannaisia, joilla käydään kauppaa sekä pörsseissä että kahdenvälisillä markkinoilla (engl. *over-the-counter market*). Optiot tarjoavat mahdollisuuksia eri suojaus- tai sijoitusstrategioihin, koska markkinoilla on tarjolla ja mahdollista kehittää monenlaisia optioita, joiden palkkiokuvauksissa vain mielikuvitus on rajana. Tässä tutkielmassa keskitytään kuitenkin vain yleisimpiin perusoptioihin, joita on olemassa kahdenlaisia: osto- ja myyntioptiot. Osto-optio (engl. *call option*) antaa määritelmän mukaan haltijalleen oikeuden, mutta ei velvollisuutta, ostaa sen kohde-etuutena oleva osake tiettyinä päivinä (erääntymispäivä, engl. *expiration date*) ennalta määrättyyn hintaan (toteutushinta, engl. *strike price* tai *exercise price*). Myyntioptiossa (engl. *put option*) puolestaan haltijalla on oikeus myydä osake vastaavin ehdoin kuin osto-optiossa. Tässä tutkielmassa käsiteltävät optiot ovat eurooppalaisia, mikä tarkoittaa, että niiden toteutus on mahdollista ainoastaan erääntymispäivänä. (Hull 2011, 7–8.)

Johdannaiset voivat olla hyvinkin monimutkaisia sijoitustuotteita. Siksi on tärkeää ymmärtää niiden palkkiokuvaus, jotta hinnoittelu on ylipäättään mahdollista. Optiot antavat vain osto- tai myyntioikeuden eivätkä pakottavaa velvollisuutta, joten niiden hinta ei voi olla negatiivinen. Jos kohde-etuuden hinta on toteutuspäivänä option kannalta epäedullinen, ei optiota kannata toteuttaa, vaan vastaavan transaktion voi tehdä edullisemmin suoraan kohde-etuudella. Optioiden palkkiokuvaus muodostuu edusta, joka syntyy kohde-etuuden ja toteutushinnan välisestä erosta. Osto- ( $C$ ) ja myyntioption ( $P$ ) palkkiot maturiteetissa  $T$  ovat

$$C(T, K) = (S_T - K)^+ = \max(S_T - K, 0) \quad (3.1)$$

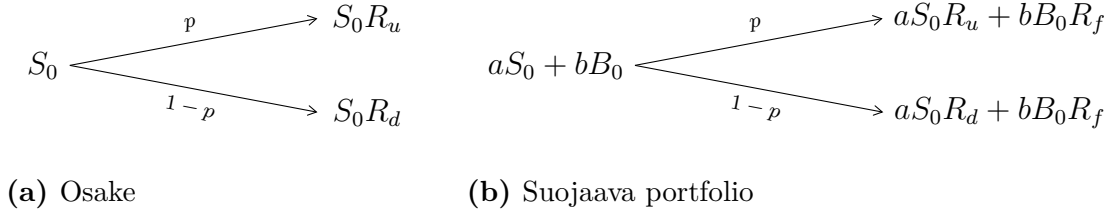
$$P(T, K) = (K - S_T)^+ = \max(K - S_T, 0) \quad (3.2)$$

missä  $K$  on option toteutushinta (Hull 2011, 198). Palkkiokuvauksista nähdään, että osto-optio on kannattavaa toteuttaa silloin, kun kohde-etuuden hinta on toteutushintaa korkeampi, ja myyntioptio sen sijaan päinvastaisessa tilanteessa. Tämä logiikka on seurausta siitä, että osto-option haltija pystyy option avulla ostamaan kohde-etuuden halvemmalla ja myyntioption haltija puolestaan myymään kalliimmalla kuin suoraan pörssin noteeraamalla senhetkiselä hinnalla.

Futuuri- ja termiinisopimuksiin sitoutuminen ei vaadi sopimuksen solmimishetkellä kummaltakaan osapuolelta rahallista maksua (Hull 2011, 6). Sen sijaan optioiden mahdollistama oikeus ostaa tai myydä kohde-etuutta ei tule ilmaiseksi, minä takia optioita on kyettävä hinnoittelemaan oikein jo ennen erääntymispäivää.

Koska tämän tutkielman empiirisessä osassa käsitellään S&P 500 -indeksiä kohde-etuutenaan käyttäviä optioita, käytetään nyt hinnoittelun soveltavissa esimerkeissä kohde-etuutena osakeindeksiä, mikä on teorian kannalta yhtäpitävää yhden yksittäisen osakkeen kanssa.

Tutustutaan aluksi yksinkertaisiin yhden periodin binomimarkkinoihin. Cox *ym.* (1979) kehittivät binomihinnoittelumallin, jossa osakemarkkinadynamiikkaa kuvastaa tila-avaruus  $\Omega = (\omega_1, \omega_2)$ , missä realisoitumistodennäköisyydet ovat objektiivisessa maailmassa  $\mathbb{P}[\omega_1] = p \in (0, 1)$  ja  $\mathbb{P}[\omega_2] = 1 - p$ . Maailmantiloihin sidottu tuottovektori on  $\mathbf{R} = (R_u, R_d)$  ja riskitön tuotto  $R_f$ , joille vaaditaan lisäksi  $0 \leq R_d < R_f < R_u$ . Kuvion 3a binomipuu kuvaa osakemarkkinan mahdollisia realisaatioita yhden periodin aikana.



**Kuvio 3:** Yhden periodin binomipuu

Seuraavaksi Cox *ym.* (1979) muodostavat portfolion osakkeesta  $S$  ja riskittömästä korkosijoituksesta (bondista)  $B$ . Koska mallin mukaan portfolio voidaan valita mielivaltaisesti, sillä voidaan suojata mikä tahansa markkinoilta löytyvä johdannainen  $\Lambda$ , jonka kohde-etuutena osake  $S$  on. Valitaan  $\Lambda$ :n suojaavaan portfolioon  $a$  kappaletta osaketta  $S$  ja  $b$  kappaletta bondia  $B$ . Osake voi tuottaa maailmantilasta riippuen joko  $R_u$  tai  $R_d$ , mutta bondin tuotto on varmasti  $R_f$  maailmantilasta riippumatta. Suojaavan portfolion tuottoa on kuvattu kuviossa 3b. Koska johdannainen  $\Lambda$  voidaan suojata täydellisesti, täytyy arbitraasittomuuden vallitessa suojaavan portfolion ja johdannaisen hinnan olla yhtä suuret sekä alkuhetkellä  $t = 0$  että maturiteetissa  $t = 1$  eli

$$\begin{cases} \Lambda_u = aS_0 R_u + bB_0 R_f \\ \Lambda_d = aS_0 R_d + bB_0 R_f \\ \Lambda_0 = aS_0 + bB_0 \end{cases} \Leftrightarrow \begin{bmatrix} R_u & R_f \\ R_d & R_f \\ 1 & 1 \end{bmatrix} \begin{bmatrix} aS_0 \\ bB_0 \end{bmatrix} = \begin{bmatrix} \Lambda_u \\ \Lambda_d \\ \Lambda_0 \end{bmatrix} \quad (3.3)$$

Tilasisidonnaisille tuotoille ja riskittömälle korolle yllä esitetty vaatimus erisuuruudesta ja järjestyksestä takaa kerroinmatriisin säännöllisyyden ja siksi mahdollistaa portfolion painokertoimien  $a$  ja  $b$  ratkaisemisen yhtälöryhmän (3.3) kahden ensimmäisen

mäisen yhtälön avulla:

$$\begin{cases} a = \frac{\Lambda_u - \Lambda_d}{(R_u - R_d)S_0} \\ b = \frac{R_u\Lambda_d - R_d\Lambda_u}{R_f B_0(R_u - R_d)} \end{cases}$$

Sijoittamalla nämä painokertoimet edelleen yhtälöryhmän kolmanteen yhtälöön se sievenee muotoon

$$\Lambda_0 = \frac{1}{R_f} [q\Lambda_u + (1 - q)\Lambda_d], \quad q = \frac{R_f - R_d}{R_u - R_d} \in (0, 1) \quad (3.4)$$

Björk (2009, 7) selventää tuottojen arbitraasittomuusehtoa intuitiivisella tulkinnalla: osakkeen tuotto ei saa dominoida riskitöntä tuottoa kaikissa maailmantiloissa tai päinvastoin. Jos tällainen tilanne olisi mahdollinen, olisi kannattavaa ottaa lyhyt positio heikommin pärjäävään kohteeseen ja sijoittaa kaikki varallisuus dominoivaan kohteeseen, mikä mahdollistaisi äärettömän arbitraasin.

Cox ym. (1979) ovat siis osoittaneet, että johdannaisen nykyhetken arvo voidaan nähdä diskontattuna odotusarvona. Nyt binomimallin siirtymätodennäköisyydet ovat kuitenkin  $q$  ja  $1 - q$ , jotka perustuvat ainoastaan maailmantiloista riippuviin tuottoihin. Alkuperäiset objektiivisen maailman todennäköisyydet ovat siis täysin merkityksettömiä arbitraasittoman hinnoittelun kannalta. Tällainen todennäköisyysrakenne muodostaa markkinoille ekvivalentin riskineutraalin martingaalimitan  $\mathbb{Q}$  ja näin ollen riskineutraali hinnoittelu on mahdollista käyttämällä pelkääntään tuottoperustaisia todennäköisyyksiä (Björk 2009, 8). Yleistämällä yhtälö (3.4) koskemaan mitä tahansa ehdollista vaadetta  $X$  riskineutraali ja arbitraasiton hinnoittelu voidaan siis kirjoittaa muodossa

$$\pi_0[X] = \frac{1}{R_f} \mathbb{E}^{\mathbb{Q}}[X], \quad (3.5)$$

missä  $\pi_0$  kuvaa nykyhintaa (Björk 2009, 12). Osto- ja myyntioptioihin sovellettuna niiden käyvät nykyhinnat voidaan ilmaista muodossa

$$\begin{aligned} C(0, K) &= \frac{1}{R_f} \mathbb{E}^{\mathbb{Q}}[(S_1 - K)^+] = \frac{1}{R_f} [q(S_0 R_u - K)^+ + (1 - q)(S_0 R_d - K)^+] \\ P(0, K) &= \frac{1}{R_f} \mathbb{E}^{\mathbb{Q}}[(K - S_1)^+] = \frac{1}{R_f} [q(K - S_0 R_u)^+ + (1 - q)(K - S_0 R_d)^+] \end{aligned}$$

Jos edellä kuvattu riskineutraali martingaalimita  $\mathbb{Q}$  on olemassa ja markkinat ovat sen seurauksena arbitraasittomat, on tarpeellista tutkia, miten objektiiviset todennäköisyydet maailmassa  $\mathbb{P}$  ja riskineutraalit todennäköisyydet maailmassa  $\mathbb{Q}$  liittyvät toisiinsa. Kyseeseen tulee stokastinen diskonttaustekijä eli deflaattori  $\mathcal{Q}$ , jonka avulla suoritetaan mitan muunnos käyttämällä Cameron–Martin–Girsanov-teoriaa.

Deflaattori määritellään eräänlaisena uskottavuuskertoimenä tai toiselta nimeltään Radon–Nikodym-derivaattana

$$\mathcal{Q} = \frac{d\mathbb{Q}}{d\mathbb{P}}$$

joka pysyy ajan suhteen vakiona ja jonka odotusarvo on aina yksi. (Björk 2009, 162–164.) Osakkeiden hintoja mallinnettaessa tällainen sopiva stokastinen deflaattori on muotoa

$$\mathcal{Q}_a = e^{-aZ - \frac{1}{2}a^2}, \quad a \in \mathbb{R}, \quad (3.6)$$

jonka parametri  $a$  on mielivaltaisesti valittava reaaliluku ja  $Z \sim N(0, 1)$  on standardia normaalijakaumaa noudattava satunnaismuuttuja.

Mitan muunnos maailmasta  $\mathbb{P}$  riskineutraaliin maailmaan  $\mathbb{Q}$  tapahtuu deflaatoimalla  $\mathbb{P}$ :n alainen odotusarvo yhtälön (3.6) mukaisella deflaattorilla. Näin ollen ehdollisen vaateen  $X$  riskineutraali odotusarvo on muotoa

$$\mathbb{E}^{\mathbb{Q}}[X] = \mathbb{E}^{\mathbb{P}}[\mathcal{Q}_a X].$$

Käyttämällä ehdollisen vaateen tilalla osakkeen  $S_t$  yhden periodin kasvutekijää yhtälöstä (2.10) saadaan kasvun odotusarvoksi mitan  $\mathbb{Q}$  alaisuudessa ratkaistua

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}[X] &= \mathbb{E}^{\mathbb{P}} \left[ e^{-aZ_t - \frac{1}{2}a^2} e^{\mu - \frac{1}{2}\sigma^2 + \sigma Z_t} \right] \\ &= e^{\mu - \frac{1}{2}\sigma^2 - \frac{1}{2}a^2} \mathbb{E}^{\mathbb{P}} \left[ e^{(\sigma - a)Z_t} \right] \\ &= e^{\mu - \frac{1}{2}\sigma^2 - \frac{1}{2}a^2} e^{\frac{1}{2}(\sigma - a)^2} \\ &= e^{\mu - \frac{1}{2}\sigma^2 - \frac{1}{2}a^2 + \frac{1}{2}\sigma^2 + \frac{1}{2}a^2 - \sigma a} = e^{\mu - \sigma a} \end{aligned}$$

Koska mitta  $\mathbb{Q}$  halutaan riskineutraaliksi, pakotetaan keskimääräinen kasvu riskittömän koron mukaiseksi eli

$$e^{\mu - \sigma a} = e^{r_f}$$

ja ratkaistaan tästä deflaattorin parametriksi

$$a^* = \frac{\mu - r_f}{\sigma}.$$

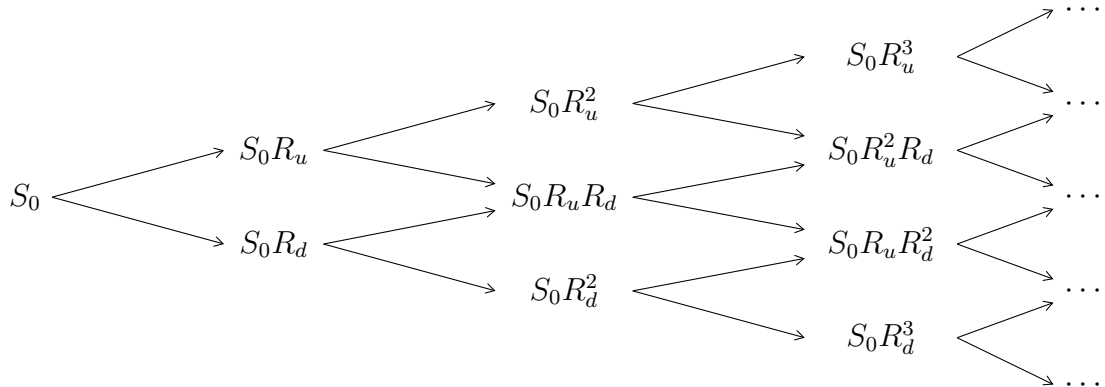
Tällä parametrivalinnalla  $a^*$  saadaan siis minkä tahansa osakkeen tuotto  $\mu$  muunnettua riskittömäksi koroksi mitan  $\mathbb{Q}$  alaisuudessa, minkä vuoksi mittaa  $\mathbb{Q}$  kutsumaan juuri riskineutraaliksi martingaalimitaksi. (Kwok 2008, 112–113.) Arbitraasittomuuden vallitessa, eli kun parametrin  $a^*$  arvo on vakio kaikille riskillisille kohde-etuuksille, mittamuunnos voidaan käytännössä suorittaa automaattisesti, jolloin voidaan suoraan hyödyntää riskitöntä korkoa kaikessa hinnoittelussa.

Mittamuunnoksella vaikutetaan vain satunnaisuutta ajavan faktorin  $Z_t$  odotusarvoon. Jos mitan  $\mathbb{P}$  alaisuudessa  $Z_t \sim N(0, 1)$ , niin silloin  $\mathcal{Q}_a$ :lla deflaatoimalla

mitan  $\mathbb{Q}$  alaisuudessa  $\tilde{Z}_t = Z_t + a \sim N(0, 1)$ . Tätä havaintoa voidaan hyödyntää yksittäisen  $Z_t$ :n lisäksi Wiener-prosessin yhteydessä, sillä myös  $\mathbb{P}$ -mitallinen Wiener-prosessi voidaan muuntaa riskineutraalin mitan  $\mathbb{Q}$  alaiseksi tavanomaiseksi Wiener-prosessiksi. On myös syytä huomata, että mittamuunnos ei vaikuta prosessin varianssiin millään tavalla. (Björk 2009, 164–165.)

### 3.2 Yleisemmät diskreetit hinnoittelumallit

On selvää, että edellisessä luvussa kuvatut markkinat eivät kuvasta todellisuutta kovin hyvin, koska tila-avaruus koostui vain kahdesta mahdollisesta maailmantilasta ja aikahorisontti käsitti vain yhden periodin eteenpäin. Lähdetään laajentamaan edellä kuvattua yhden periodin binomimallia ajassa pidemmäksi kuvion 4 mukaisesti.



**Kuvio 4:** Useamman periodin binomipuu

Jokaisella periodilla yhdestä solmusta voidaan edetä kahteen suuntaan binomimallin periaatteen mukaisesti: realisoituu joko tuotto  $R_u$  tai  $R_d$ , kuten yhden periodin mallissakin. Toisin sanoen useamman periodin binomipuu koostuu sarjasta vastaavia yhden periodin binomimalleja, joiden logiikkaa on kuvattu kuviossa 3a. Jos osakkeen hinta on tiedossa hetkellä  $t$ , voidaan maturiteettiin ulottuva osakkeen hintaprosessi ilmaista muodossa

$$S_T = S_t R_u^j R_d^{T-t-j} \quad \text{todennäköisyydellä} \quad \binom{T-t}{j} p^j (1-p)^{T-t-j}, \quad (3.7)$$

missä  $j \sim \text{Bin}(T-t, p) \in [0, T-t]$  on binomijakautunut satunnaismuuttuja, joka kuvaa maailmantilan  $\omega_1$  realisaatioiden määrää aikavälillä  $[t, T]$ . (Capinski ja Zastawniak 2003, 56–57.)

Jos arbitraasittomuusehto  $R_d < R_f < R_u$  on voimassa, se takaa automaattisesti riskineutraalin martingaalimitan olemassaolon vastaavasti kuin yhden periodin tapauksessakin. Tällöin hinnoittelussa voidaan siirtyä suoraan yhtälön (3.5) kuvaaman  $\mathbb{Q}$ -maailman odotusarvon laskemiseen todennäköisyyksillä  $q$  ja  $1-q$ , jotka on

määritelty yhtälön (3.4) yhteydessä. Useamman periodin binomimallin tapauksessa johdannaisen  $\Lambda(S_T)$  käypä arvo hetkellä  $t$  on siis

$$\begin{aligned}\pi_t[\Lambda] &= \frac{1}{R_f^{T-t}} \mathbb{E}^{\mathbb{Q}}[\Lambda(S_T)] \\ &= \frac{1}{R_f^{T-t}} \sum_{j=0}^{T-t} \left[ \binom{T-t}{j} q^j (1-q)^{T-t-j} \Lambda(S_t R_u^j R_d^{T-t-j}) \right].\end{aligned}\quad (3.8)$$

(Björk 2009, 24.) Jos kyseessä on osto-optio, on kaavassa (3.8) oleva ehdollinen vaade  $\Lambda$  tällöin muodoltaan palkkiokuvauksen (3.1) mukaisesti  $\Lambda(\cdot) = (S_t R_u^j R_d^{T-t-j} - K)^+$  ja myyntioption tapauksessa puolestaan palkkion (3.2) mukaisesti  $\Lambda(\cdot) = (K - S_t R_u^j R_d^{T-t-j})^+$ . Kaavassa (3.8) summataan siis kaikki mahdolliset maturiteetissa tapahtuvat palkkiot, joita painotetaan tilakohtaisilla binomitodennäköisyyksillä.

Binomimalleissa rajoitutaan vain kahden maailmantilan toteutumiseen. Vaikka periodien jakoa voidaan tihentää eli yhden periodin pituutta lyhentää esimerkiksi yhdestä päivästä minuuttiin tarkkuuden ja realistisuuden lisäämiseksi, voi rajallinen tapausavaruus olla silti liian raju yksinkertaistus reaalimaailmasta. Laajennetaan seuraavaksi tapausavaruutta niin, että maailmantiloja on enemmän kuin kaksi:  $\boldsymbol{\Omega} = (\omega_1, \omega_2, \dots, \omega_M)_{1 \times M}$ . Samalla voidaan tarkastella useampaa kohdeetuutta, joiden transponoitu hintavektori on  $\mathbf{S}_t^T = (S_t^1, S_t^2, \dots, S_t^N)_{1 \times N}$ , mutta rajoitutaan kuitenkin aluksi yhden periodin markkinoihin  $t = 0, 1$ . Yhdistämällä hintavektori tapausavaruuteen saadaan muodostettua hintamatriisi

$$\mathbf{D}_t = \begin{bmatrix} S_t^1(\omega_1) & S_t^1(\omega_2) & \cdots & S_t^1(\omega_M) \\ S_t^2(\omega_1) & S_t^2(\omega_2) & \cdots & S_t^2(\omega_M) \\ \vdots & \vdots & \ddots & \vdots \\ S_t^N(\omega_1) & S_t^N(\omega_2) & \cdots & S_t^N(\omega_M) \end{bmatrix}_{N \times M} \quad t = 0, 1 \quad (3.9)$$

Lähdetään tutkimaan yllä kuvattujen markkinoiden arbitraasittomuutta muodostamalla allokaatiovektori  $\mathbf{h} \in \mathbb{R}^N$ , joka kuvaa kunkin markkinoilla olevan kohdeetuuden painoa. On selvää, että portfolion arvoprosessi on muotoa

$$V_t^{\mathbf{h}} = \mathbf{h} \mathbf{D}_t = \sum_{i=1}^N h^i S_t^i \quad (3.10)$$

$$V_t^{\mathbf{h}}(\omega_j) = \mathbf{h} \mathbf{D}_t(\omega_j), \quad \forall j = 1, \dots, M$$

molemmilla ajanhetkillä  $t = 0, 1$ . Arbitraasiportfolioksi kutsutaan sellaista portfolioa, jolla voidaan saavuttaa positiivista tuottoa ilman tappioriskiä ja nettoinvestointia. Täsmällisemmin arbitraasiportfolion ehdot ovat siis

$$\begin{aligned}V_0^{\mathbf{h}} &= 0 \\ \mathbb{P}[V_1^{\mathbf{h}} \geq 0] &= 1 \\ \mathbb{P}[V_1^{\mathbf{h}} > 0] &> 0\end{aligned}\quad (3.11)$$

missä ensimmäinen ehto vaatii portfolion  $\mathbf{h}$  muodostamista niin, ettei ulkopuolista rahoitusta tarvita, toinen ehto poistaa tappioriskin kaikista maailmantiloista  $\omega_j, j = 1, \dots, M$  ja viimeisen ehdon mukaan on olemassa todellinen mahdollisuus positiiviseen tuottoon vähintään yhdessä maailmantilassa. (Björk 2009, 27.)

Edellä kuvatun arbitraasiportfolion olemassaoloa voidaan tutkia Farkasin lemmän avulla. Kwokin (2008, 41, 43) sovelluksen mukaan arbitraasiportfoliota ei voi olla olemassa vain, jos on olemassa ei-negatiivinen vektori  $\boldsymbol{\psi} \in \mathbb{R}_+^M$  niin, että

$$\mathbf{S}_0 = \mathbf{D}_1 \boldsymbol{\psi}. \quad (3.12)$$

Björkin (2009, 28–32) sovelluksessa puolestaan  $\mathbf{D}_1$  on skaalattu matriisiin ensimmäisellä rivivektorilla (kohde-etuudella), jolloin on kyseessä suhteellinen hintamatriisi, jonka numeräärinä on ensimmäinen kohde-etuus. Koska  $\mathbf{D}_1$  on suhteellinen ja vapaasti valittavissa, on yhtälön ratkaiseva vektori  $\boldsymbol{\psi}$  siksi Arrow–Debreu-arvopaperien (ks. Arrow ja Debreu 1954) eli niin sanottujen alkeistila-arvopaperien hintavektori (tilahintavektori).

Tilahinnat  $\boldsymbol{\psi}$  muodostavat eri maailmantilojen realisoitumistodennäköisyydet, koska ne hinnoittelevat saman kiinteän palkkion jokaisessa eri tilassa. Oikeastaan voidaan puhua myös maailmantilojen diskreetistä tiheysfunktioista. Todennäköisyyskäsite saadaan muodostettua skaalaamalla tilahinnat niiden summalla eli

$$q_j = \frac{\psi_j}{\sum_{j=1}^M \psi_j}, \quad \forall j \quad \Leftrightarrow \quad \mathbf{q} = \frac{1}{\mathbb{1}^\top \boldsymbol{\psi}} \boldsymbol{\psi}, \quad (3.13)$$

koska jokainen  $\boldsymbol{\psi}$ :n komponentti on arbitraasittomuusehdon vuoksi positiivinen ja skaalauksen takia  $q_j$ :t summautuvat yhdeksi. Jos muodostetaan kaikista alkeistilarvopapereista portfolio yhtäläisellä painotuksella, saavutetaan varmasti täysin riskitön tuotto. Näin ollen alkeistilahintojen summa vastaa riskittömän sijoituksen hintaa, mistä saadaan johdettua riskitön tuotto

$$\mathbb{1}^\top \boldsymbol{\psi} = \frac{1}{R_f}.$$

Käyttämällä hinnoittelussa arbitraasittomuuden takaavia todennäköisyyksiä  $\mathbf{q}$  ja riskitöntä korkoa  $R_f$  puhutaan ekvivalentista riskineutraalista martingaalimitasta  $\mathbb{Q}$  vastaavasti kuin aiemmin luvussa 3.1. Siksi myös aiemmin esitetty yleinen hinnoitteluperiaate (3.5) toimii useamman maailmantilan tapauksessa. (Björk 2009.)

Ei ole kuitenkaan aina itsestäänselvää, että ekvivalentti martingaalimita  $\mathbb{Q}$  on olemassa tai varsinkaan yksikäsitteinen. Voi olla, että markkinoilla olevien arvopaperien määrä tai riippuvuus toisistaan vaikuttaa hintamatriisin  $\mathbf{D}$  asteeseen niin, että yksikäsitteistä ehdon  $\boldsymbol{\psi} > 0$  toteuttavaa ratkaisua yhtälöön (3.12) ei ole. Jos ratkaisua ei ole ollenkaan, eivät markkinat ole arbitraasittomat. Jos puolestaan ratkaisuja on useampi, on markkinoilla olemassa useampia eri arbitraasittomuuden takaavia martingaalimittoja eli arbitraasittomuuden takaavia hintoja. (Björk 2009, 37.)

Markkinoiden täydellisyys tarkoittaa määritelmän mukaisesti, että mille tahansa mielivaltaiselle ehdollistetulle vaateelle  $X$  on olemassa lineaarikombinaatio  $\mathbf{h}$  markkinoilla olevista arvopapereista  $\mathbf{D}$ , joilla vaade saadaan täysin suojattua (Björk 2009, 35). Välttämätön ehto markkinoiden täydellisyydelle on, että markkinoilla on olemassa vähintään yhtä monta kohde-etuutta kuin maailmantiloja. Tämä ei kuitenkaan välttämättä riitä, koska jos kohde-etuudet ovat toisistaan riippuvia, ei vaateen suojaava allokaatio ole yksikäsitteinen. Siksi yksikäsitteinen ja riittävä ehto markkinoiden täydellisyydelle on, että hintamatriisin  $\mathbf{D}$  asteen tulee olla maailmantilojen määrä  $M$ . (Kwok 2008, 50–51.) Toisin sanoen markkinoilla tarvitaan vähintään  $M$  kappaletta toisistaan lineaarisesti riippumatonta kohde-etuutta.

### 3.3 Optiohinnoittelu normaalijakaumaa käyttäen

Aiemmissa luvuissa edettiin yksinkertaisesta binomimallista jo hieman edistyneempään usean maailmantilan malliin. Voidaan kuitenkin kysyä, kuvaavatko diskreetit mallit riittävän hyvin markkinoiden todellista käyttäytymistä. Luvussa 2.1 esitettiin, että osakemarkkinoiden tuottojen voidaan olettaa noudattavan normaalijakaumaa, joten on perusteltua mallintaa ehdollisten vaateiden hinnoittelua normaalijakauman avulla. Aloitetaan hinnoittelun tutkiminen yhden periodin markkinoista ja siirrytään sen jälkeen jatkuva-aikaiseen maailmaan.

Ajautuvan satunnaiskulkuprosessin (2.2) perusteella osakkeen hinnan muutokset ovat normaalisti jakautuneita. Yhtälön (3.5) kuvaamaa yleistä hinnoittelusääntöä riskineutraalissa maailmassa voidaan soveltaa myös jatkuvaan jakaumaan, mikäli markkinoiden tuottorakenteen perusteella on olemassa riskineutraali martingaalimitta  $\mathbb{Q}$ . Soveltamalla yhtälön (2.10) mukaista osakkeen hintadynamiikkaa yhtälöiden (3.1) ja (3.2) mukaisiin optioiden palkkioihin odotusarvot ovat

$$\begin{aligned} C(0, K) &= \frac{1}{R_f} \mathbb{E}^{\mathbb{Q}}[(S_T - K)^+] \\ &= \frac{1}{R_f} \mathbb{E}^{\mathbb{Q}}\left[\left(S_0 e^{r_f - \frac{1}{2}\sigma^2 + \sigma\tilde{Z}} - K\right)^+\right] \\ P(0, K) &= \frac{1}{R_f} \mathbb{E}^{\mathbb{Q}}[(K - S_T)^+] \\ &= \frac{1}{R_f} \mathbb{E}^{\mathbb{Q}}\left[\left(K - S_0 e^{r_f - \frac{1}{2}\sigma^2 + \sigma\tilde{Z}}\right)^+\right] \end{aligned}$$

Osto- ja myyntioptiot tuottavat palkkion, kun

$$\begin{aligned} \tilde{Z} &\geq -\frac{\ln\left(\frac{S_0}{K}\right) + r_f - \frac{1}{2}\sigma^2}{\sigma} := -L \quad (\text{osto-optio}) \\ \tilde{Z} &\leq -L \quad (\text{myyntioptio}) \end{aligned}$$

ja käyttämällä näitä epäyhtälöitä optioiden hinnat saadaan muotoon

$$\begin{aligned} C(0, K) &= \frac{1}{R_f} \mathbb{E}^{\mathbb{Q}} \left[ \left( S_0 e^{r_f - \frac{1}{2}\sigma^2 + \sigma \tilde{Z}} - K \right); \tilde{Z} \geq -L \right] \\ &= S_0 \mathbb{E}^{\mathbb{Q}} \left[ e^{\sigma \tilde{Z} - \frac{1}{2}\sigma^2}; \tilde{Z} \geq -L \right] - K e^{-r_f} \mathbb{Q} \left[ \tilde{Z} \geq -L \right] \\ P(0, K) &= \frac{1}{R_f} \mathbb{E}^{\mathbb{Q}} \left[ \left( K - S_0 e^{r_f - \frac{1}{2}\sigma^2 + \sigma \tilde{Z}} \right); \tilde{Z} \leq -L \right] \\ &= K e^{-r_f} \mathbb{Q} \left[ \tilde{Z} \leq -L \right] - S_0 \mathbb{E}^{\mathbb{Q}} \left[ e^{\sigma \tilde{Z} - \frac{1}{2}\sigma^2}; \tilde{Z} \leq -L \right] \end{aligned}$$

Nyt odotusarvojen sisällä on molempien optioiden tapauksessa yhtälön (3.6) mukainen deflaattori  $\mathcal{Q}_{-\sigma}$ . Odotusarvo saadaan helposti ratkaistua, kun otetaan käyttöön ekvivalentti mitta  $\mathbb{M}$ , jonka alaisuudessa deflaattorin odotusarvo onkin nyt rajaavan välin realisoitumistodennäköisyys: osto-option tapauksessa  $\mathbb{M}[\tilde{Z} \geq -L]$  ja myyntioption tapauksessa  $\mathbb{M}[\tilde{Z} \leq -L]$ . Mitan  $\mathbb{M}$  alaisuudessa uusi normaalisti jakautunut satunnaismuuttuja on ajautuman kautta yhteydessä mitan  $\mathbb{Q}$  alaiseen satunnaismuuttujaan  $\hat{Z} = \tilde{Z} - \sigma \sim N(0, 1)$  ja tätä tietoa hyödyntämällä optioiden hinnat voidaan sieventää muotoon

$$C(0, K) = S_0 \Phi(L + \sigma) - K e^{-r_f} \Phi(L) \quad (3.14)$$

$$P(0, K) = K e^{-r_f} \Phi(-L) - S_0 \Phi(-L - \sigma) \quad (3.15)$$

missä  $\Phi(\cdot)$  on standardin normaalijakauman kertymäfunktio.

Edellä kuvatut malliyhtälöt (3.14) ja (3.15) soveltuvat sellaisenaan vain yhden periodin markkinoille. Olisi kuitenkin hyödyllisempää pystyä hinnoittelemaan optioita minä tahansa hetkenä millä tahansa juoksuaajalla. Siksi yleistetään edellä olevaa mallia jatkuva-aikaiseen kehikkoon. Käytännössä tämä tarkoittaa, että satunnaisuutta ajava normaalisti jakautunut satunnaismuuttuja  $Z$  vaihdetaan määritelmän (2.4) mukaiseksi Wiener-prosessiksi. Näin ollen osakkeen hintaprosessina käytetään suoraan yhtälön (2.10) kuvaamaa dynamiikkaa. Riskineutraalit odotusarvot eli optioiden käyvät arvot hetkellä  $t < T$  ovat siis esitettävissä muodossa

$$\begin{aligned} C(t, K) &= e^{-r_f} \mathbb{E}_{(t, S_t)}^{\mathbb{Q}} \left[ (S_T - K)^+ \right] \\ &= e^{-r_f} \mathbb{E}_{(t, S_t)}^{\mathbb{Q}} \left[ \left( S_t e^{(r_f - \frac{1}{2}\sigma^2)(T-t) + \sigma(\tilde{W}_T - \tilde{W}_t)} - K \right)^+ \right] \\ P(t, K) &= e^{-r_f} \mathbb{E}_{(t, S_t)}^{\mathbb{Q}} \left[ (K - S_T)^+ \right] \\ &= e^{-r_f} \mathbb{E}_{(t, S_t)}^{\mathbb{Q}} \left[ \left( K - S_t e^{(r_f - \frac{1}{2}\sigma^2)(T-t) + \sigma(\tilde{W}_T - \tilde{W}_t)} \right)^+ \right] \end{aligned}$$

missä  $\tilde{W}$  on Wiener-prosessi  $\mathbb{Q}$ -mitan alaisuudessa. On syytä huomata, että Wiener-prosessin muutos  $\tilde{W}_T - \tilde{W}_t \sim \sqrt{T-t} \tilde{Z} \sim N(0, T-t)$  on samoin jakautunut kuin tavanomainen normaalisti jakautunut satunnaismuuttuja, jonka varianssina on aikavälin pituus. Tätä havaintoa hyödyntäen optiot tuottavat palkkion vastaavalla

tavalla kuin yhden periodin tapauksessa, eli kun

$$\tilde{Z} \geq -\frac{\ln\left(\frac{S_t}{K}\right) + \left(r_f - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} := -L(t, K) \quad (\text{osto-optio}) \quad (3.16)$$

$$\tilde{Z} \leq -L(t, K) \quad (\text{myynti-optio}) \quad (3.17)$$

Jatkuva-aikaisessa tapauksessa edetään samoin kuin yhden periodin mallissakin, mutta nyt huomioiden aikavälin pituus, jolloin optioiden käyviksi arvoiksi saadaan

$$C(t, K) = S_t \Phi\left(L(t, K) + \sigma\sqrt{T-t}\right) - Ke^{-r_f(T-t)} \Phi(L(t, K)) \quad (3.18)$$

$$P(t, K) = Ke^{-r_f(T-t)} \Phi(-L(t, K)) - S_t \Phi\left(-L(t, K) - \sigma\sqrt{T-t}\right) \quad (3.19)$$

Yhtälöt (3.18) ja (3.19) kuvaavat juuri sitä optiohinnoittelun mullistanutta mallia, jonka Black ja Scholes (1973) sekä Merton (1973) kehittivät. Sen ansiosta kaupan käynti optioilla yleistyi huomasti, koska markkinaosapuolilla oli käytössään tehokas työkalu, jonka avulla johdannaisia saattoi hinnoitella oikein.

Edellä kuvatut malliyhtälöt eivät sisällä osinkotuottoa, vaan soveltuvat ainoastaan osinkoa maksamattomille kohde-etuuksille. Tässä tutkielmassa kohde-etuutena hyödynnetään S&P 500 -indeksiä, johon kuuluu osinkoa maksavia yhtiöitä. Kuten riskitön korko Black–Scholes–Merton-kehikossa on vakio, myös osinkotuoton kohdalla oletetaan sen olevan deterministinen vakio  $d$ . Osingon irrotessa juuri ennen hetkeä  $t$  osakkeen hinta laskee osingon määrän hetkeen  $t$  siirryttäessä. Siten osakkeen hintaprosessin (2.6) ajautumatermi korjataan jatkuvalla osinkotuotolla niin, että osakkeen hintaa kuvaava geometrinen Brownin liike on muotoa

$$dS_t = (\mu - d)S_t dt + \sigma S_t dW_t. \quad (3.20)$$

Hintaprosessiin tehty muutos näkyy myös epäyhtälöissä (3.16) ja (3.17) niin, että nyt vakioisella jatkuvalla osinkotuotolla

$$L_d(t, K) = \frac{\ln\left(\frac{S_t}{K}\right) + \left(r_f - d - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}$$

ja samoin Black–Scholes–Merton-optiohinnoitteluyhtälöt (3.18) ja (3.19) tulevat muotoon

$$C(t, K) = S_t e^{-d(T-t)} \Phi\left(L_d(t, K) + \sigma\sqrt{T-t}\right) - Ke^{-r_f(T-t)} \Phi(L_d(t, K)) \quad (3.21)$$

$$P(t, K) = Ke^{-r_f(T-t)} \Phi(-L_d(t, K)) - S_t e^{-d(T-t)} \Phi\left(-L_d(t, K) - \sigma\sqrt{T-t}\right) \quad (3.22)$$

(Björk 2009, 230, 239; Kwok 2008, 129–130.)

### 3.4 Black–Scholes–Merton-mallin vääristymät oikeilla markkinoilla

Corrado ja Su (1996) tiedostavat edellä kuvatun Black–Scholes–Merton-kaavan hyödyllisyyden optioiden hinnoittelussa, mutta samalla toteavat sen sisältävän joitakin virheellisyyksiä, jotka johtuvat mallin melko tiukoista oletuksista liittyen kohde-etuuden taustalla vaikuttavaan hintaprosessiin. Black ja Scholes (1972, 400)<sup>3</sup> itse listaavat kehittämäänsä malliin liittyvät oletukset:

1. kohde-etuuden volatilitteetti  $\sigma$  on vakio option juoksuaikana
2. riskitön korko  $r_f$  on vakio juoksuaikana ja sekä otto- että antolainauksen korko on sama
3. option haltija on suojautunut mahdollisen voitonjaon vaikutuksilta kohde-etuuden hintaan
4. kohde-etuusosakkeen tuotto on log-normaalisti jakautunut, mikä implikoi hintaprosessin noudattavan geometrista Brownin liikettä
5. muut yleiset markkinoiden toimintaan liittyvät ehdot, jotka muun muassa jättävät kaupankäyntikustannukset huomioimatta, takaavat sijoittajille loputtoman lyhyeksimyymämahdollisuuden ja mahdollisuuden lainata riskittömällä korolla rajattomasti.

Lisäksi he huomauttavat, että malli ei siinä mielessä ole kovin rajoittunut, koska siinä ei kuitenkaan tarvita markkinoiden yhteistä näkemystä eri kohde-etuuksien odotetuista tuotoista tai niiden yhteisjakaumasta. Myöskään kohde-etuuksien käyppiin arvoihin malli ei ota kantaa. Nämä lievennykset johtuvat siitä, että mallin pohjalla on oletus arbitraasittomuudesta ja sitä kautta riskineutraalin martingaalimitan olemassaolosta, jolloin odotettuna tuottona voidaan käyttää riskitöntä korkoa. Kohde-etuuden hinnoittelussa mahdollisesti esiintyvä virheellisyys käyvän arvon näkökulmasta ei myöskään vaikuta option arvoon, koska optio on johdannainen, jonka hinta ei määräydy kohde-etuuden käyvän arvon, vaan suoraan sen hinnan mukaan – olkoon hinta kuinka väärässä tai oikeassa tahansa.

Ensimmäinen Black–Scholes–Merton-mallin epävarmuustekijä liittyy siihen, että mallissa käytettävistä parametreista yhtä, volatilitteettia, ei voida tarkistaa markkinoilta havaintohetkellä, vaan se on estimoitava. Kirjallisuudessa esiintyy useita eri näkemyksiä volatilitteetin estimointiin, mutta alan standardiksi on vakiintunut historiallisen otosvarianssin neliöjuuri. (Ncube ja Satchell 1997, 289.) Toiseksi on syytä ainakin epäillä, ovatko markkinat aina arbitraasittomat. Mikäli arbitraasimahdollisuuksia on olemassa, se estää riskineutraalin martingaalimitan olemassaolon, minkä

<sup>3</sup> Selvennettäköön lukijalle, että tämä artikkeli (Black ja Scholes 1972) on kirjoitettu myöhemmin kuin artikkeli, jossa optiohinnoittelumalli esiteltiin (Black ja Scholes 1973). Hinnoittelumallin esittelevän artikkelin (1973) julkaisussa kesti useampia vuosia, minkä vuoksi artikkelien julkaisuvedot eivät tässä yhteydessä vastaa kirjoitusjärjestystä.

vuoksi markkinoiden valtavassa kokonaisuudessa ei ole olemassa yksiselitteistä käsitystä tuottodynamiikasta. Kolmas virheellisyyksiä aiheuttava tekijä liittyy parametrien deterministisyyteen.

Corrado ja Su (1996) puuttuvat volatilitietin vakioisuusoletukseen ja toteavat selkeästi, että osaketuottojen volatilitiettikin on stokastinen prosessi. Varianssimallit alkoivat yleistyä, kun Engle (1982) kehitti autoregressiivisen ehdollisen heteroskedastisuuden mallin (engl. *autoregressive conditional heteroskedasticity*, ARCH), jonka Bollerslev (1986) yleistyi muutamaa vuotta myöhemmin GARCH-malliksi (engl. *generalized autoregressive conditional heteroskedasticity*), josta on ajan saatossa tullut yksi yleisimmistä varianssin mallintamiseen käytetyistä malleista. Stokastisen volatilitietin mallien kirjo alkoi kasvaa entisestään, kun muiden muassa Hull ja White (1987) sekä Heston (1993) kehittivät alan kirjallisuudessa kohtuullisen suosituiksi nousseet mallinsa. Hull–White-malli lähtee siitä, että hintadynamiikkaa kuvataan kahdella stokastisella differentiaaliyhtälöllä

$$\begin{aligned}dS_t &= \mu S_t dt + \sigma_t S_t dW_t^1 \\d\sigma_t^2 &= \alpha \sigma_t dt + \beta \sigma_t dW_t^2\end{aligned}$$

missä  $\sigma_t^2$  on kohde-etuuden ajasta riippuva varianssi sekä  $W^1$  ja  $W^2$  ovat keskenään korreloituneita Wiener-prosesseja. Varianssi noudattaa siis omaa geometrista Brownin liikettä, mikä on stokastisen volatilitietin perusajatus. Heston-mallissa hintaprosessia kuvaava differentiaaliyhtälö on sama kuin edellä, mutta volatilitiettia kuvaava stokastinen prosessi sen sijaan on muotoa

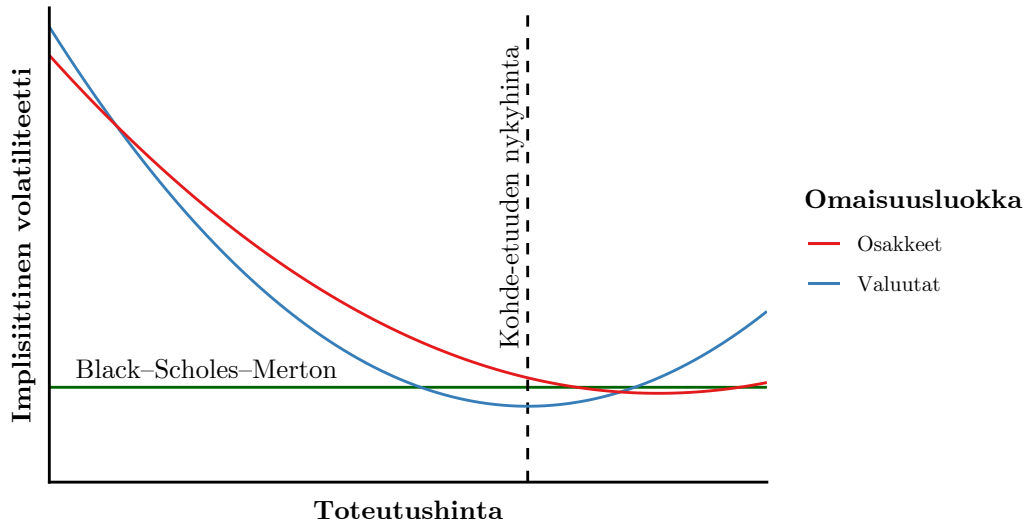
$$d\sigma_t = -\alpha \sigma_t dt + \delta dW_t^2,$$

joka noudattaa Ornstein–Uhlenbeck-prosessia. Se voidaan ratkaista Itô–Döblinin lauseen avulla muotoon

$$d\sigma_t^2 = \kappa(\theta - \sigma_t^2) dt + \beta \sigma_t dW_t^2.$$

Samoin kuin Hull–White-mallissa, myös Hestonin mallissa Wiener-prosessit  $W^1$  ja  $W^2$  ovat korreloituneet keskenään.

Edellä esitetyt poikkeavuudet ovat syitä, miksi oletus log-normaalijakautuneista tuotoista ei välttämättä toimi oikeilla markkinoilla ja aiheuttavat muun muassa aiemmin luvussa 2.3 käsiteltyä pitkää muistia osakkeiden tuottoihin. Kun todellinen jakauma ei vastaa enää Black–Scholes–Merton-mallin oletusta, se alkaa hinnoitella väärin erityisesti optioita, joiden toteutushinta on kaukana kohde-etuuden nykyisestä hinnasta. Väärinhinnoittelua voidaan tutkia volatilitietihymyn avulla. Siinä toteutushintaa lukuun ottamatta samanlaisten optioiden implisiittinen volatilitietti kuvataan toteutushinnan funktiona. Esimerkkejä volatilitietihymyistä on kuviossa 5.



**Kuvio 5:** Volatiliteettihymyjä eri omaisuusluokissa (mukaillen Hull 2011, 411, 414)

Jos Black-Scholes-Merton-malli pitäisi paikkansa, tulisi volatiliteettihymyn olla suora vaakaviiva (piirretty vihreällä), koska mallin tulisi hinnoitella kaikki optiot oikein samalla volatiliteetillä riippumatta toteutushinnasta. Useimmiten tilanne oikeilla markkinoilla on kuitenkin kuvion 5 muiden käyrien kaltainen: lähellä kohde-etuuden nykyistä hintaa markkinoilta havaittava hinta on lähellä Black-Scholes-Merton-mallin mukaista hintaa, mutta mitä kauemmas siirrytään, sitä enemmän Black-Scholes-Merton-malli alihinnoittelee optioita. Valuuttamarkkinoilla volatiliteettihymy on yleisesti melko symmetrinen kohde-etuuden hinnan muodostaman pystyakselin ympärillä, mutta konveksi. Osakemarkkinoilla volatiliteettihymyllä puolestaan on laskeva trendi, mutta myöskin se on konveksi. Epäsymmetrisyyden takia osakemarkkinoiden yhteydessä puhutaankin joskus myös volatiliteettivinoudesta (engl. *volatility skew*). Yksi tulkinta vinouteen on, että yrityksen arvon laskiessa sen oman pääoman osuus vähenee, minkä vuoksi sen riski kasvaa. (Hull 2011.)

Volatiliteettihymyllä tai -vinoudella on yhteys implisiittiseen todennäköisyysjakaumaan. Shimko (1993, 35) selventää, että volatiliteettihymyn konveksisuus tarkoittaa huipukasta ja siten paksuhäntäisempää jakaumaa. Samalla tavalla laskeva trendi volatiliteettihymyssä on yhteydessä jakauman negatiiviseen vinouteen, jolloin todennäköisyysmassaa on enemmän vasemmassa hännässä. Molemmat ominaisuudet ovatkin osakemarkkinoilla empiirisesti havaittuja ja yleisesti tunnettuja tuloksia (ks. esim. Fama 1965; Mills 1995; Corrado ja Su 1996).

### 3.5 Optioiden sisältämä informaatio tulevaisuudesta

Ross (1976) tarkastelee markkinoiden täydellisyttä tila-avaruuden ollessa lähes jatkuva. Koska mahdollisia maailmantiloja on enemmän kuin arvopapereita, ei markkinoilla ole mahdollisuutta saavuttaa mielivaltaista kassavirtaa. Havainnollistetaan tätä yksinkertaisen esimerkin avulla: markkinoilla on vain yksi osake, jonka maailmantiloihin  $(\omega_1, \omega_2, \omega_3)$  sidotut hinnat seuraavalla periodilla ovat

$$\mathbf{D}_{t+1} = \begin{bmatrix} 60 & 100 & 150 \end{bmatrix}.$$

Havaitaan, että koska hintamatriisin  $\mathbf{D}$  aste on vain 1, ei markkinoilla oleva yksi arvopaperi riitä suojaamaan mielivaltaista kassavirtaa. Jos markkinoille kuitenkin kehitetään kaupankäynnin kohteeksi kaksi tätä osaketta kohde-etuutenaan käyttävää osto-optiota toteutushinnoilla  $K_1 = 80$  ja  $K_2 = 120$ , tulee hintamatriisi muotoon

$$\mathbf{D}_{t+1} = \begin{bmatrix} 60 & 100 & 150 \\ 0 & 20 & 70 \\ 0 & 0 & 30 \end{bmatrix}, \quad |\mathbf{D}_{t+1}| \neq 0,$$

jolloin saadaan muodostettua portfolio, jolla on mahdollista replikoida mikä tahansa mielivaltainen kassavirta, koska hintamatriisi on säännöllinen. Useimmiten johdannaisten laskeminen liikkeelle on markkinoiden kannalta tehokkaampaa ja halvempaa kuin toisen osakkeen liikkeellelasku (Ross 1976, 76).

Tärkeä huomio Rossin (1976) tutkimuksesta on se, että optioilla, tai johdannaisilla ylipäänsä, voidaan täydentää markkinoita, jotta markkinoiden täydellisyys voidaan saavuttaa vähemmällä määrällä perinteisiä arvopapereita. Jatkuvatilaisessa maailmassa täydellisyys ei voida päästä tällä tavoin huomatta äärettömän monta johdannaista. Kasvattamalla kaupankäynnin kohteena olevien arvopapereiden määrää erilaisten johdannaisten avulla saadaan kuitenkin kattavampi arvopapereiden verkosto, joka kuvaa tilajatkumoa tarkemmin, vaikkakin edelleen diskreetisti.

Breeden ja Litzenberger (1978) hyödyntävät yllä kuvattua havaintoa muodostaessaan alkeistila-arvopapereita pelkkien osto-optioiden avulla. Alkeistila-arvopaperilla (myös Arrow–Debreu-arvopaperi)  $\mathbf{e}_\omega$  tarkoitetaan arvopaperia, jonka palkkio maturiteetissa  $T$  on euklidinen yksikkövektori, jonka  $\omega$ :s alkio on 1:

$$\mathbf{e}_\omega = \underbrace{(0, \dots, 1, \dots, 0)}_{M \text{ kpl}}$$

Breeden ja Litzenberger (1978) olettavat aluksi, että kohde-etuuden  $S_T$  todennäköisyysjakauma on diskreetti ja realisaatiot tasavälisiä, ja luovat osto-optioita jokaiselle mahdolliselle  $S_T$ :n tilalle ( $M$  kpl). Yksinkertaisuuden vuoksi oletetaan, että kohde-etuuden hinta  $S_T$  eri tiloissa ja optioiden toteutushinnat  $K$  ovat luonnollisia lukuja

alkaen yhdestä. Luotujen osto-optioiden palkkioita eli käyviä arvoja maturiteetissa on kuvattu taulukossa 2.

**Taulukko 2:** Osto-optioiden palkkioita kohde-etuuden tilan mukaan

$S_T$	$C(T, 0)$	$C(T, 1)$	$C(T, 2)$	$\dots$	$C(T, M - 1)$
1	1	0	0	$\dots$	0
2	2	1	0	$\dots$	0
3	3	2	1	$\dots$	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$M$	$M$	$M - 1$	$M - 2$	$\dots$	1

Koska taulukossa 2 toteutushintajako on tasavälinen ja välin jako on yksi rahyksikkö, siirryttäessä korkeamman toteutushinnan optioon eli yksi sarake oikealle siirtyy sarakkeen arvot yhden rivin alaspäin, kunnes viimeisessä sarakkeessa optio tuottaa yhden rahyksikön palkkion vain viimeisessä maailmantilassa. Jos nyt jokaisen option palkkiosta vähennetään yhtä toteutushintaa suuremman option palkkio, saadaan aikaan alakolmiomatriisi, jonka alakolmio on täynnä ykkösiä (kaavan (3.23) ensimmäinen erotus). Tehtäessä sama vähennys vielä uudestaan, saavutetaan tilanne, jossa optioiden palkkioita kuvaava matriisi onkin identiteettimatriisi (kaavan (3.23) jälkimmäinen erotus). Nythän kyseessä on joukko alkeistilarvopapereita, koska toteutushinnan  $K$  option palkkio on tasan yksi vain siinä maailmantilassa, jossa toteutushinta vastaa kohde-etuuden hintaa  $K = S_T$ . Toisin sanoen alkeistilarvopaperi  $e_K$  voidaan replikoida portfoliolla

$$[C(T, K_{i-1}) - C(T, K_i)] - [C(T, K_i) - C(T, K_{i+1})] \quad (3.23)$$

$$= C(T, K_{i-1}) - 2C(T, K_i) + C(T, K_{i+1}). \quad (3.24)$$

Jokaisen alkeistilarvopaperin käypä hinta on siis saatavilla optioiden hintojen perusteella millä tahansa ajanhetkellä  $t \leq T$ , kunhan markkinoiden arbitraasittomuusehdot täyttyvät. (Breedon ja Litzenberger 1978, 626.)

Yleisemmin, jos kohde-etuuden hinnan ja samalla toteutushintojen jako maailmantilojen välillä on  $\Delta_\omega S$ , saadaan alkeistilarvopaperin  $e_\omega$  käypä arvo  $\psi$  määritettyä yhtälöstä

$$\psi(S_T, T; \Delta_\omega S) = \frac{1}{\Delta_\omega S} [C(T, S_T - \Delta_\omega S) - 2C(T, S_T) + C(T, S_T + \Delta_\omega S)]$$

Jakamalla yhtälö nyt puolittain välin pituudella  $\Delta_\omega S$  ja antamalla välin pituuden lähestyä nollaa saadaan raja-arvo

$$\lim_{\Delta_\omega S \rightarrow 0} \frac{\psi(S, T; \Delta_\omega S)}{\Delta_\omega S} = \left. \frac{\partial^2 C(T, K)}{\partial K^2} \right|_{K=S_T} \quad (3.25)$$

joka on option hinnan toinen derivaatta toteutushinnan suhteen. Kohde-etuuden hinnan  $S_T$  noudattaessa jatkuvaa jakaumaa yksittäisen realisaation todennäköisyys on periaatteessa nolla. Kohde-etuuden hinnalla  $S_T$  on tässä tapauksessa kuitenkin tiheysfunktio, joka saadaan yhtälön (3.25) perusteella option hintafunktiosta  $C(T, K)$  olettaen, että  $C$  on kahdesti differentioituva. On huomionarvoista, ettei minkäänlaisia oletusta kohde-etuuden satunnaisuutta ajavasta dynamiikasta ole tehty. (Breedon ja Litzenberger 1978, 625–628.) Markkinoilla noteerattavien optioiden hintojen perusteella saadaan siis määritettyä tiheysfunktio, jonka pohjalla ei tarvita mitään oletusta markkinoiden hintaprosesseista, vaan kaikki tarvittava informaatio on optioiden hinnoissa. Black–Scholes–Merton-mallin hinnoitteluyhtälöissä (3.21) ja (3.22) ainut parametri, jota ei saada markkinoilta suoraan, on volatilitteetti, joka on samalla ainut niin sanotusti vapaa parametri. Siksi luvussa 3.4 esitelty volatilitteettihymy sisältää optioiden hintojen perusteella informaatiota tulevista odotuksista, joiden perusteella markkinoilla vallitseva implisiittinen tiheysfunktio voidaan muodostaa.

Implisiittisen todennäköisyysjakauman estimointi optiohintojen avulla on ollut kirjallisuudessa melko suosittua ja monet tutkijat ovatkin kehittäneet omia versioitaan estimointimenetelmästä. Pääasiallinen jako menetelmien välillä kulkee jakauman parametrisuudessa tai parametrittomuudessa. Parametrisissa menetelmissä ei ole löydetty yleisesti hyväksi havaittua muotoa, minkä takia menetelmässä on suuri riski spesifioida malli jo lähtökohtaisesti väärin. Tämän vuoksi ei-parametriset menetelmät tarjoavat yleisen käsityksen mukaan luotettavampia estimaatteja, vaikka vaarana onkin, että jakauma tulee helposti monihuippuiseksi ja vääristyneeksi mallin ylisovittamisen takia. (Taboga 2016, 454.)

## 4 AINEISTON JA SOVELLETUN TUTKIMUSMENETELMÄN KUVAUS

### 4.1 Aineiston hankinta ja alustavat muokkaustoimenpiteet

Taulukossa 3 on eritelty aineistotyypeittäin tutkielman empiirisessä osassa käytetyt aikasarja-aineistot. Alkuvaiheessa on käytetty testiotosta, joka haettiin syksyllä 2019. Varsinaiset aineistot, joita tutkielmassa lopulta hyödynnetään, on noudettu 17.3.2020.

**Taulukko 3:** Tutkielmassa käytetyt aikasarja-aineistot

Aineistotyyppi	Haetut aikasarjat
Aikaväli Lähde	
<b>S&amp;P 500 -indeksi</b> 3.1.2000–16.3.2020 Refinitiv Eikon (Datastream-lisäosa)	Hintaindeksi Kokonaistuottoindeksi
<b>Indeksioptiot</b> 20.6.2018–16.3.2020 Refinitiv Eikon (Datastream-lisäosa)	Kaikki osto- ja myyntioptiot: Erääntymispäivä Toteutushinta Markkinahinta Kaupankäyntivolyymi
<b>Riskitön korko</b> 3.1.2000–16.3.2020 Federal Reserve Bank of St. Louis (2020)	1 kk 3 kk 6 kk 1 v 2 v 3 v 5 v

Eikon-ohjelmistosta haetut eli osakeindeksi- ja optioaikasarjat ovat yhdessä Excel-tiedostossa. Yhdysvaltojen keskuspankin aineistotietokannasta ([Federal Reserve Bank of St. Louis 2020](#)) haetut korkoaikasarjat ovat .csv-muotoisissa tekstitiedostoissa jokainen maturiteetti erikseen. Aineistolle suoritetaan joitakin välttämättömiä muokkaustoimenpiteitä, jotta se saadaan työstettävään muotoon.

**S&P 500 -indeksi:** Osakemarkkinoita kuvaavan S&P 500 -indeksin tiedot vietään alkuperäisestä Excel-tiedostosta omaksi .csv-tiedostoksi aineistonhallinnalli-

sista syistä ja R:ssä `.xlsx`-tiedostojen tuonnissa ilmenneiden epävakauksien<sup>4</sup> johdosta. `.csv`-tiedostosta aineisto on tuotu R:ään omaksi objektiksi.

**Korkodata:** Riskittömät korot tuodaan omista `.csv`-tiedostoistaan R:ään, minkä jälkeen ne yhdistetään yhdeksi objektiksi. Näin aineisto saadaan helposti käsiteltävään muotoon, jossa jokaisen päivän kohdalla kunkin maturiteetin havainnot ovat omilla sarakkeissaan.

**Indeksiotiot:** Eri optioita tähän tutkielmaan kuuluvassa aineistossa on tuhansia. Excelin rajoitteista johtuen alkuperäinen aineisto on transponoituna niin, että päivät ovat sarakkeissa ja eri optiot ja niiden aikasarjat puolestaan riveillä. Excelin Visual Basic -ohjelmointikielellä (VBA) aineisto järjestetään uudelleen niin, että päivät etenevät riveittäin ja sarakkeissa on käytettävien aikasarjojen havainnot. Apusarakkeena käytetään option yksilöivää itse luotua kaupankäyntitunnusta. Excelissä yhden välilehden rivimäärä on rajoitettu 1 048 576:een, joten aineisto täytyy tässä vaiheessa lisäksi jakaa neljälle välilehdelle. Kaikki välilehdet vietään omiksi `.csv`-tiedostoiksi<sup>4</sup>, jotka tuodaan R:ään. Päivähavainnot on tässä vaiheessa yhteensä 3 721 202 kappaletta.

R:ssä muodostetaan kaavamaisesti kaupankäyntitunnus kaikille optioille niiden tyyppin, erääntymispäivän ja toteutushinnan mukaan: `<Tyyppi>_<Erääntymispäivä>_<Toteutushinta>`. Sen jälkeen suodatetaan aineistosta pois sellaiset optiot, joilla ei ole olemassa hintaa, ei ole käyty kauppaa päivän aikana tai juoksuaikaa ei ole enää jäljellä. Suodatuksessa havainnoista lähtee pois suurin osa ja jäljelle jää 306 194 havaintoa.

Optio-objektia käytetään työskentelyssä eniten, joten sen jokaiselle riville yhdistetään havaintopäivää vastaava S&P 500 -indeksin taso, lasketaan jäljellä oleva juoksuaika vuosina ja sen perusteella valitaan maturiteetiltaan sopivin riskitön korko. Lisäksi option toteutushinta muunnetaan suhteelliseksi senhetkiseen indeksin tasoon nähden  $K/S_t - 1$ , jotta säilytetään vertailukelpoisuus eri päivien välillä.

## 4.2 Menetelmän perusteet

Breeden ja Litzenberger (1978) keksivät tavan tutkia tilahintoja ja johtaa niiden perusteella markkinoiden implikoima todennäköisyysjakauma optioiden avulla. Ensimmäinen menetelmän soveltaja oli Shimko (1993), jonka sovellusta tässä tutkielmasa pääasiassa seurataan. Aihe on tosin ollut suosittu tutkimuskohde, minkä vuoksi menetelmistä on ajan mittaan kehitetty tutkijoiden omia versioita. Pyrin tässä tutkielmassa yhdistämään parhaimpia puolia olemassa olevasta kirjallisuudesta.

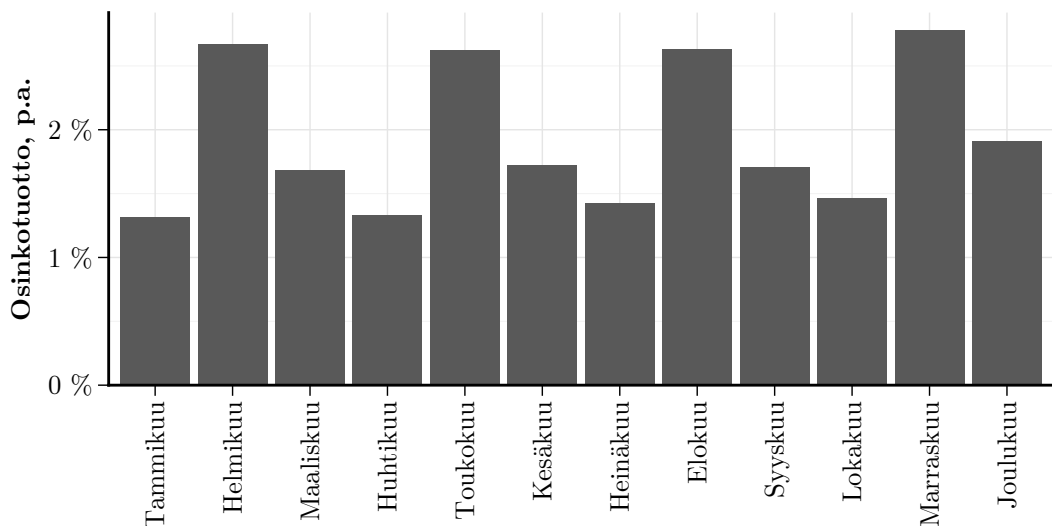
<sup>4</sup> Aineistosta hävisi tietoja, kun tiedosto tuotiin `.xlsx`-muodossa R:ään. `.csv`-muodossa tätä ongelmaa ei esiintynyt.

Olen suorittanut analyysin käyttämällä avoimen lähdekoodin R-ohjelmistoa (R Core Team 2020) ja RStudio-käyttöliittymää. Dataintensiivisyyden vuoksi pitkäketoisimpia laskutoimituksia siirryttiin ajamaan pilvipalvelimella yön yli. Liitteessä A on kuvattu käytetyt R-paketit ja ohjelmistot tarkemmin. Seuraavaksi esittelen analyysin etenemisen vaihe vaiheelta.

### 4.3 Analyysin alkutoimenpiteet

**Osinkotuoton estimointi:** Koska käytän tutkielmassani kohde-etuutena S&P 500 -indeksiä, johon kuuluvat yhtiöt maksavat osinkoa, on osinkotuotto otettava huomioon, koska maksettavat osingot laskevat indeksin pistelukua. Osinkotuottoa ei ollut saatavilla suoraan tietokannoista, joten se on estimoitu seuraavalla tavalla: indeksin kokonaistuottoversion prosentuaalisesta tuotosta on vähennetty indeksin pistelukuun perustuva prosentuaalinen tuotto, mikä kuvastaa päivittäistä osinkotuottoa.

Seuraavaksi on oletettu, että indeksiin kuuluvat yhtiöt maksavat osinkoa tasaisesti. Osinkotuotto on siten laskettu kuukauden tasolla summaamalla päivittäiset havainnot. Kuukausittaisista osinkotuotoista on laskettu keskiarvot 2000-luvulta ja lopulta prosentuaalinen tuotto on muunnettu vuosittaiseksi logaritmiseksi tuotoksi. Kuukausittainen osinkotuoton jakautuminen vuoden ajalle on havainnollistettu kuviossa 6.



**Kuvio 6:** S&P 500 -indeksin annualisoidun osinkotuoton jakautuminen eri kuukausille

Kuvion perusteella havaitaan, että on tärkeää ottaa huomioon osinkotuoton vaihtelu kuukauden mukaan, koska option jäljellä olevalle juoksuajalle pitäisi käyt-

tää juuri sen aikaista osinkotuottoa oikean hinnan määrittämiseksi. Tästä syystä jokaiselle havaintopäivälle on kaikille käytetyille optioille estimoitu osinkotuotto jäljellä olevan juoksuajan painotettuna keskiarvona. Tämä ottaa juoksuajan osumisen eri kuukausille hyvin huomioon, mikä ei onnistuisi vakiolla tai vuositason osinkotuotolla. Painotuksessa on käytetty viisipäiväistä kaupankäyntiviikkoa ja operaation suoraviivaisen automaation takia juhlapyyhiä tai muita päiviä, jolloin pörssit ovat olleet kiinni, ei ole otettu huomioon.

On vielä kuitenkin huomautettava, että indeksin pisteluvut on ilmoitettu vain kahden desimaalin tarkkuudella, joten numeerista epätarkkuutta esiintyi jonkin verran päivittäisten osinkotuottojen laskennassa. Tutkielman kannalta riittävään tarkkuuteen kuitenkin päästään aggregoimalla osinkotuotto kuukausitasolle ja lisäksi käyttämällä pitkän aikavälin keskiarvoa, joka yhtyy kohtalaisen hyvin esimerkiksi Damodaranin (2020) tuloksiin.

**Arbitraasittomuusehtojen toteutuminen:** Kuten esimerkiksi Jackwerth ja Rubinstein (1996), myös tässä tutkielmassa poistetaan havaintoaineistosta sellaiset optiot, jotka rikkovat arbitraasittomuusehtoja. Arbitraasillisten optioiden olemassaolo myös estäisi riskineutraalin todennäköisyysmitan olemassaolon, mikä olisi hyvin ongelmallista tutkielman tavoitteiden kannalta. Osto- ja myyntioptioiden hintojen ylä- ja alarajat ovat

$$S_t e^{-d\Delta t} - K e^{-r_f \Delta t} \leq C \leq S_t e^{-d\Delta t}$$

$$K e^{-r_f \Delta t} - S_t e^{-d\Delta t} \leq P \leq K e^{-r_f \Delta t}$$

(Hull 2011, 350) ja havainnot, jotka eivät täytä epäyhtälöiden kuvaamia rajoitteita, yksiselitteisesti poistetaan aineistosta. Hinnan yläraja ei rikkoudu yhdessäkään havainnossa, mutta alarajalla joudutaan poistamaan 7 025 havaintoa, joista osto-optioita on 4 054 ja myyntioptioita 2 971 kappaletta. Lisäksi tässä vaiheessa tulee poistetuksi sellaiset havainnot, joille ei ole saatavilla riskitöntä korkoa aineistossa olevien puutteiden vuoksi. Lopulta havaintoja poistetaan yhteensä 8 679, jolloin jäljelle jää 297 515 havaintoa.

**Myyntioptiot muunnetaan osto-optioiksi:** Optioiden osto-myynti-pariteetin avulla myyntioptioista tehdään osto-optioita, jolloin tulee selkeämmäksi käsitellä vain pelkkiä osto-optioita analyysin seuraavissa vaiheissa. Myyntioption perusteella laskettava osto-option hinta on osto-myynti-pariteetin mukaan

$$C = P + S_t e^{-d\Delta t} - K e^{-r_f \Delta t},$$

kun optiot ovat toteutushinnaltaan  $K$  ja erääntymispäivältään  $T$  toisiaan vastaavat.

**Aineiston rajaaminen:** Tutkielman toteutuksen kannalta aineiston valintaan liittyy merkittävä ongelma, kun halutulla markkina-alueella ei optioilla käyty kauppaa riittävän paljon luotettavien tulosten saamiseksi. S&P 500 -indeksillä kaupankäyntiaktiivisuus on suurimmaksi osaksi korkea. Kaupankäyntiaktiivisuuden rajaksi asetetaan vähintään 100 kappaletta päivässä, minkä alle jäävät päivähavainnot joudutaan poistamaan tehokkaan hinnanmuodostuksen epävarmuudesta johtuen.

Lisäksi analyysin myöhemmässä vaiheessa vaaditaan kullekin päivälle riittävän paljon optioita, joilla on sama erääntymispäivä. Tähän rajaksi on asetettu vähintään 10 eri toteutushintaa, ja mikäli optioita on liian vähän, ne poistetaan kyseiseltä päivältä hyödyttöminä. Tämä vaihe on aineiston suhteen kaikkein vaativin, minkä vuoksi käytetystä aineistosta karsiutuu suuri osa pois. Havaintoja on tämän suodatuksen jälkeen 110 846 kappaletta eli vain noin 37 % edelliseen vaiheeseen verrattuna.

#### 4.4 Implisiittisen volatilitietin määrittäminen ja tasoitus

**Volatilitietihymy:** Black–Scholes–Merton-kaavan (3.18) kaikki muut parametrit paitsi volatilitietti ovat saatavilla markkinoilta. Koska volatilitietti on ainut vapaa parametri, voidaan se ratkaista kullekin optiohavainnolle käyttämällä markkinoilta saatavia arvoja muille parametreille. Normaalijakauman kertymäfunktioista johtuen suljetun muodon ratkaisua ei ole saatavissa, mutta numeerisesti ongelmaa lähdetään ratkaisemaan seuraavasti:

$$|C_\sigma - C| < \varepsilon,$$

missä  $C_\sigma$  on optimoitavan volatilitietin mukainen osto-option hinta,  $C$  markkinoilla noteerattu osto-option hinta ja  $\varepsilon \approx 0,000122$  on liukuluvuilla tapahtuvan numeerisen ratkaisun tarkkuuden määrittävä parametri, jonka  $R$  määrittää automaattisesti. Ratkaisualgoritmillemme annetaan alustavaksi hakuväliksi  $\sigma \in [0,0001, 1]$ , mutta tarvittaessa yläraja saa joustaa ylöspäin, mikäli volatilitietti on korkea, jolloin algoritmin nopeus hidastuu välin kasvaessa. Löydetty juuri tallennetaan implisiittisenä volatilitiettinä kullekin havainnolle.

**Implisiittisen volatilitietin tasoitus luonnollisella kuutio-spliniillä:** Grith, Härdle ja Schienle (2012) kuvailevat kolme tapaa, miten implisiittisen todennäköisyysjakauman ei-parametrisessa määrittämisessä voidaan lähteä liikkeelle. Yksi tavoista on sovittaa volatilitiettihymyyn eli implisiittisen volatilitietin funktioon toteutushinnan mukaan kahdesti differentioituva funktio. Aiheen ensimmäisiin soveltajiin lukeutuva Shimko (1993) käyttää paraabelia, mutta matala-asteisten polynomien sijaan tällä vuosituhannella on siirrytty enenevässä määrin monimutkaisempiin menetelmiin. Useissa tutkimuksissa on käytetty kuutio-spliniä (ks. esim. Campa, Chang

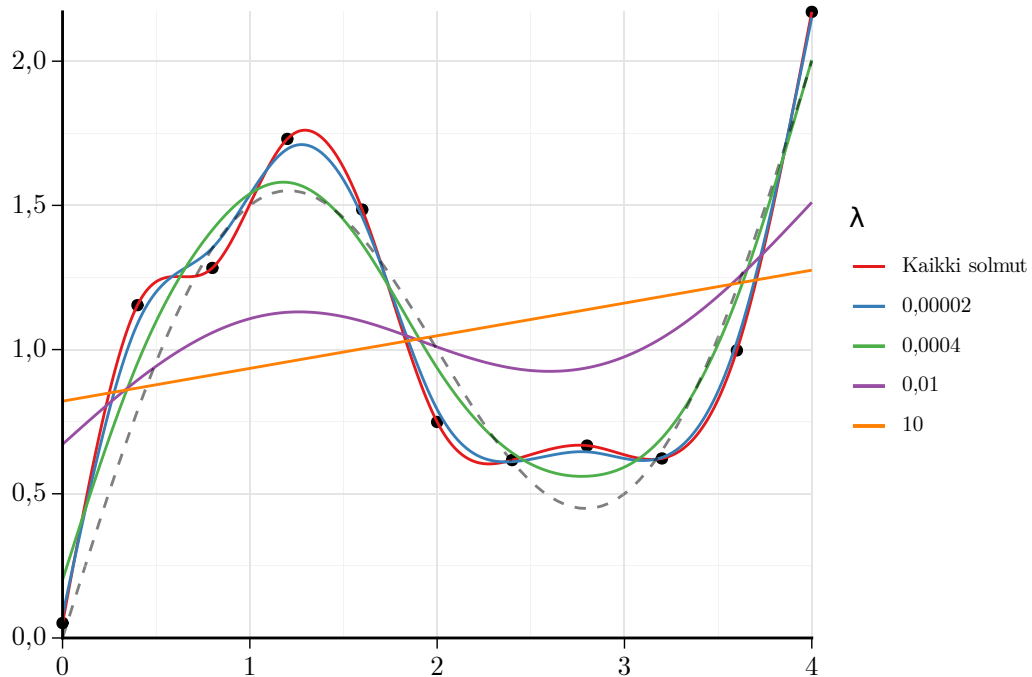
ja Reider 1998; Bu ja Hadri 2007; Monteiro, Tütüncü ja Vicente 2008; Santos ja Guerra 2015; Salazar Celis, Liang, Lemmens, Tempère ja Cuyt 2015), ja erityisesti Bliss ja Panigirtzoglou (2002) näyttävät sen toimivan paremmin kuin estimoitaessa sekoitejakauma, jossa yhdistetään kaksi log-normaalia jakaumaa.

Kuutiosplini on paloittain määritelty kolmannen asteen polynomifunktio, jossa solmut määräävät palojen reunakohdat. Kuutiosplinin jatkuvuuden takia funktion on saatava solmuissa sama arvo lähestyittäessä solmua kummasta suunnasta tahansa. Differentioituvuuden takia myös ensimmäisen ja toisen derivaatan on oltava jatkuvia. Vapaita parametreja on kuitenkin vielä enemmän kuin rajoitteita. Luonnollisessa kuutiosplinissä (engl. *natural cubic spline*) rajoitteiksi lisätään vielä, että toisten derivaattojen tulee olla nollia ensimmäisen ja viimeisen solmun kohdalla, jolloin funktio on lineaarinen reunimmaisten solmujen ulkopuolella. (Tsay ja Chen 2019, 134–138.) Solmujen määrää ei ole rajoitettu, joten täydellisen soviteen saa varmasti valitsemalla solmuiksi kaikki havaintopisteet. Interpoloimalla soviteen mukaista funktiota saataisiin mielivaltaisen tiheällä jaolla tarkka approksimaatio havaintopisteiden taustalla olevasta funktiosta.

Näin tarkka sovite ei kuitenkaan ole tämän tutkielman kannalta mielekäs tapa, koska approksimaatiofunktio tulisi ylisovitettu ja se pyrkisi selittämään myös kohinaa. Siksi olen valinnut approksimointimenetelmäksi tasoitetun splinin (engl. *smoothed spline*), jossa solmujen määrää on rajattu kohtuullisen paljon ja jossa tasoitusparametri (engl. *smoothing parameter*)  $\lambda \geq 0$  kontrolloi soviteen tasapainoa tarkkuuden ja tasaisuuden välillä (Harezlak, Ruppert ja Wand 2018, 19). Tsay ja Chen (2019, 141–142) havainnollistavat  $\lambda$ :n vaikutusta jäännösneliösummayhtälöllä

$$\text{RSS}(f, \lambda) = \sum_{i=1}^N [y_i - f(x_i)]^2 + \lambda \int f''(s) ds, \quad (4.1)$$

missä joukko  $\{x_i, y_i\}_{i=1}^N$  kuvaa havaintoarvopareja ja  $f(x)$  on mikä tahansa kahdesti differentioituva funktio. Ensimmäinen termi on tavanomainen jäännösneliösumma, jota minimoimalla saavutetaan paras mahdollinen sovite, mutta samalla toinen termi rankaisee funktion  $f(x)$  kaarevuutta.  $\lambda$ :n ollessa 0 kaarevuudesta ei rangaista ja mikä tahansa  $f(x)$ , joka sopii aineistoon, kelpaa. Jos  $\lambda = \infty$ , toista derivaattaa ei sallita ollenkaan ja näin ollen tasoitettu splini konvergoituu pienimmän neliösumman mukaiseksi regressiosuoraksi. Sopivan  $\lambda$ :n löytäminen jätetään optimoinnin kohteeksi ja sen vaikutusta on havainnollistettu kuviossa 7.



**Kuvio 7:** Luonnollinen kuutiosplini eri tasoitusparametrin arvoilla

Kuviossa 7 taustalla oleva funktio on  $f(x) = (1/2)x + \sin[(\pi/2)x]$ , johon on lisätty normaalisti jakautunutta kohinaa ( $\sigma = 0,2$ ) ja joka on kuvattu katkoviivalla. Punaisessa käyrässä kaikkia havaintopisteitä on käytetty splinin solmuina, jolloin käyrä kulkee kaikkien havaintopisteiden kautta. Muilla väreillä piirretyissä käyrissä tasoitusparametri  $\lambda$  vaihtelee pienestä (sininen) suureen (oranssi). Huomataan, että hyvin pienellä  $\lambda$ :lla käyrä ottaa liikaa huomioon kohinaa (sininen), ja liian suurella  $\lambda$ :lla funktion todellinen muoto ei tule riittävästi esille (violetti). Erittäin korkealla  $\lambda$ :lla käyrä muistuttaa jo lähes suoraa (oranssi). Graafisesti tarkasteltuna sopivin  $\lambda$  olisi vihreällä käyrällä, koska se muistuttaa kohtalaisen hyvin taustalla olevaa funktiota.

R:ssä tasoitettu splini sovitetaan toteutushinnan ja implisiittisen volatilitteen muodostamalle tasolle jokaiselle päivälle ja erääntymispäivälle `smooth.spline`-funktioilla. Sen parametreissa valitaan, että kaikkia havaintopisteitä ei käytetä solmuina, ja tasoitusparametrin haussa käytetään yleistettyä ristiinvalidointia (engl. *generalized cross-validation*, GCV). Solmujen määrää rajoitetaan toisistaan eroavien toteutushintojen määrän  $\hat{K}$  mukaan

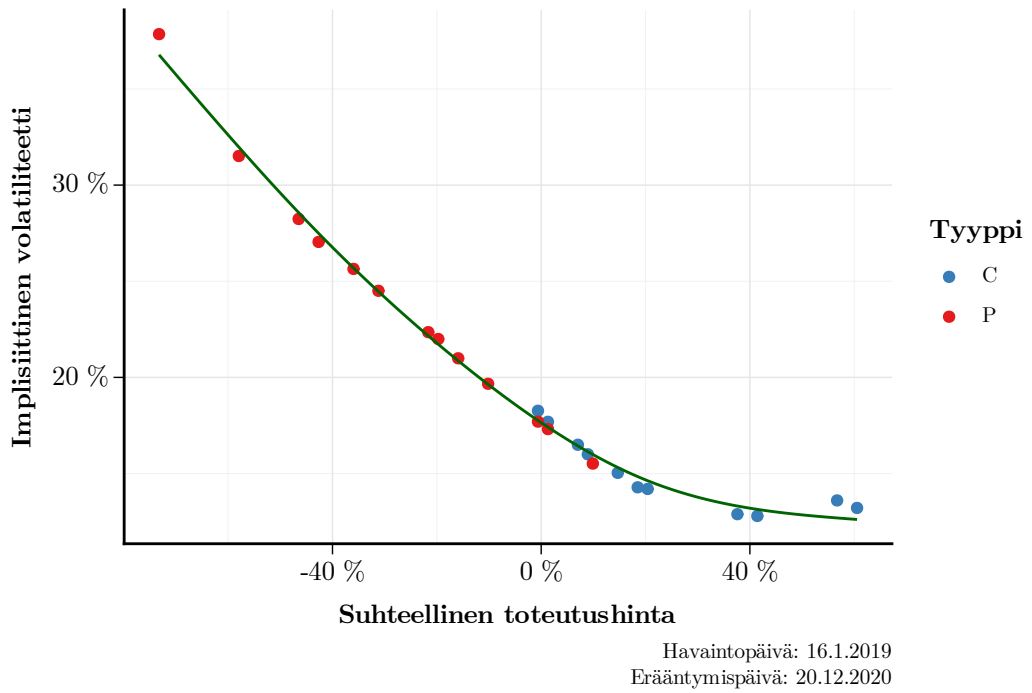
$$\kappa = 2 + \left\lceil \sqrt{\hat{K}} \right\rceil$$

missä  $\lceil x \rceil$  on pyöristysfunktio ylöspäin seuraavaan kokonaislukuun. Lisäämällä 2 taataan, että solmuja on aina vähintään 6, koska toisistaan eroavien toteutushintojen

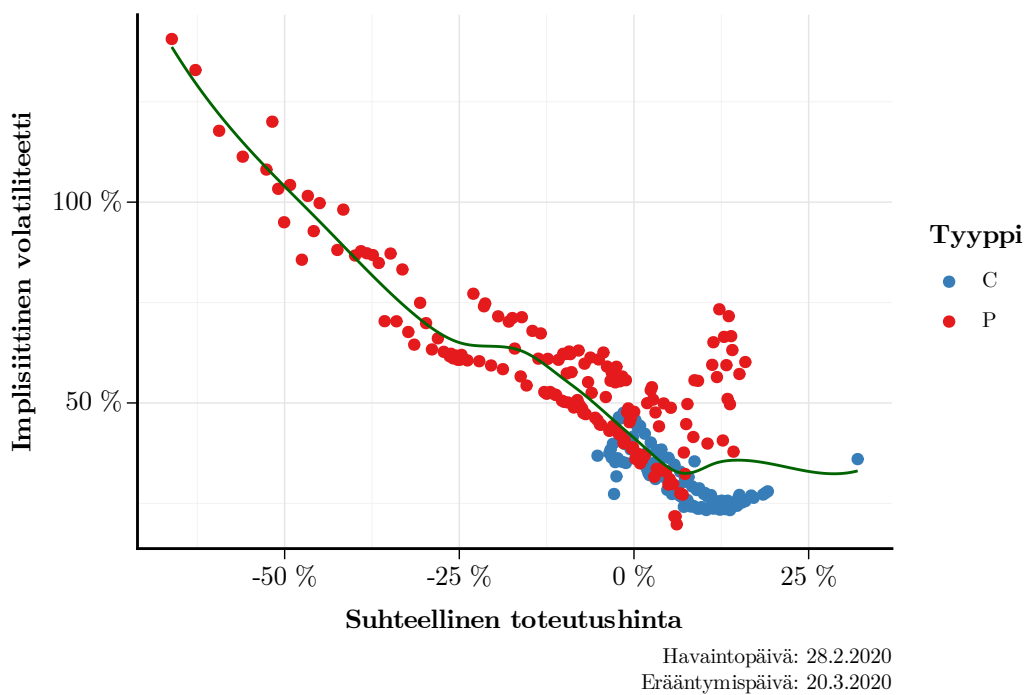
määräksi rajoitettiin aiemmin 10, jotta sovite ottaa riittävästi huomioon volatilitiiteetihymyn muodon. Neliöjuurifunktiolla rajoitetaan solmujen enimmäismäärän kasvamista liian suureksi, jotta vältetään ylisovittamiselta, mutta kuitenkin hyödynnetään tarkempaa toteutushintajakoa. Näillä rajoitteilla solmujen määrä on käytössä olevalla aineistolla  $\kappa \in [6, 20]$ .

Edellä mainitulla menetelmällä sovite näyttää graafisesti tarkasteltuna hyvältä useimpien päivien ja erääntymispäivien kohdalla. Kuitenkin joukossa on monia, etenkin jos juoksuaikaa on jäljellä enää vähän, sellaisia sovitteita, jotka ovat selvästi liian ylisovitettuja. `smooth.spline`-funktion palauttamia `spar`-arvoja (suhteellinen tasoitusparametri, josta lasketaan edellä kaavassa (4.1) esiintynyt  $\lambda$ ) lähemmin tarkasteltaessa voidaan huomata, että ylisovittamista on tapahtunut, kun `spar` on estimoitu matalaksi tai jopa negatiiviseksi. `smooth.spline`-funktio antaa mahdollisuuden rajoittaa `spar`-parametria, ja graafisesti arvioituna riittävä tasointu saadaan, kun asetetaan rajoitteeksi `spar`  $\geq 0,4$ .

Kun tasoitettua spliniä on estimoitu kaikille olemassa oleville päivien ja erääntymispäivien yhdistelmille, luodaan keinotekoisesti soviteen mukaiset implisiittiset volatilitiiteetit tasajakaiselle toteutushinta-akselille. Toteutushinnat jaetaan yhden välein alkaen kunakin päivänä pienimmästä aineistossa olevasta toteutushinnasta ja päättyen korkeimpaan toteutushintaan. Toisin sanoen implisiittinen volatilitiiteetti-ovite ainoastaan interpoloidaan, koska ekstrapoloinnissa aineisto ei enää tue approksimointia lokaalisti, minkä vuoksi pakotettu lineaarisuus saattaa vääristää häntä. Toteutushintajakoa tihentämällä havaintomäärä kasvaa huomattavan paljon, mutta sovitettu splini tuottaa kahtena päivänä ei-positiivisia volatilitiiteetteja, minkä vuoksi nämä sovitteet joudutaan poistamaan. Lopulta päivähavainnot on yhteensä 4874304 kappaletta. Kuviossa 8 on piirretty kahdelta päivästä eri tavalla onnistuneet sovitteet: osassa (a) aineisto on jo valmiiksi melko vähän vaihtelevaa, minkä vuoksi sovite istuu hyvin havaintopisteisiin, mutta osassa (b) sovite ei ole toivotun tasainen, koska aineistossa implisiittiset volatilitiiteetit ovat jo lähtökohtaisesti enemmän hajonneita eikä algoritmin mukainen tasointu siksi toimi.



(a) Riittävä tasoitus



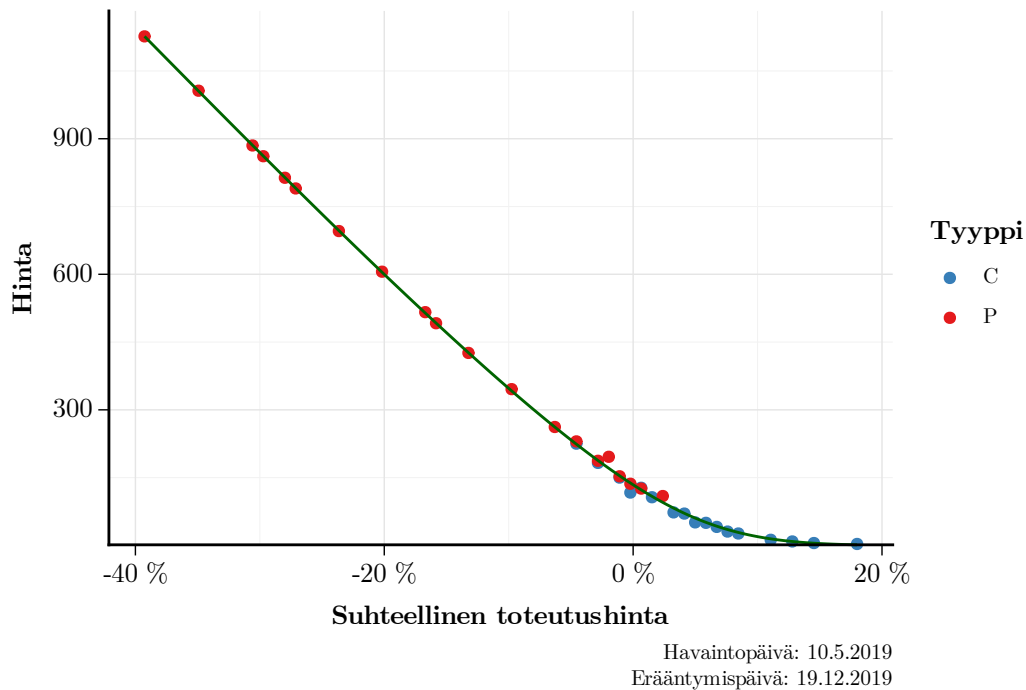
(b) Aineiston hajonta aiheuttaa riittämättömän tasoituksen

**Kuvio 8:** Implisiittisen volatilitteetin tasoitus kuutio splineillä

#### 4.5 Implisiittisen volatilitietin hyödyntäminen hintojen tasoituksessa

Vaikka volatilitietihymykuvaajista käy selkeästi ilmi, ettei Black–Scholes–Merton-mallin oletus vakiovolatilitietista pidä käytettävässä aineistossa, Bliss ja Panigirtzoglou (2002, 384–385) huomauttavat, että mallia voi silti käyttää siirtymissä eri tasojen välillä. Toisin sanoen volatilitietihymystä eli toteutushinta–volatilitietitasosta voidaan siirtyä takaisin toteutushinta–hinta-tasoon ongelmitta, vaikka taustalla oleva hintaprosessi ei todellisuudessa noudatakaan mallin mukaisia log-normaaleja lisäyksiä.

Edellä tasoitettua implisiittistä volatilitietia hyödynnetään seuraavaksi optioiden hintojen laskennassa. Uudet sovitetut hinnat lasketaan samalla toteutushinta- jaolla kuin mihin implisiittinen volatilitietti sovitettiin. Kuviossa 9 on kuvattuna erään päivän havainnot yhden erääntymispäivän optioiden hinnoista.



**Kuvio 9:** Implisiittisen volatilitietin sovitteen mukaiset optiohinnat

Sovitettu käyrä tasoittaa varsin hyvin paikoitellen poikkeavat hinnat ja muodostaa konveksin kuvan. On syytä huomioida, että kuvaajan  $X$ -akselilla on suhteellinen toteutushinta eli osto-option palkkio pienenee siirryttäessä oikealle, minkä takia käyrä on laskeva.

## 4.6 Implisiittisen todennäköisyysjakauman estimointi

Kuten luvussa 3.5 esitettiin, implisiittinen riskineutraali todennäköisyysjakauma saadaan optioiden hintojen toisesta derivaatasta toteutushinnan suhteen. Optioiden hintafunktio toteutushinnan suhteen on kuvattu myös kuviossa 9. Kaikissa tapauksissa käyrän sovitus ei onnistu yhtä hyvin kuin edellä, joten lisäksi on syytä nostaa esiin muutamia huomautuksia: Kuten yhtälöstä (3.1) voidaan päätellä, osto-option palkkio on monotonisesti ei-kasvava toteutushinnan suhteen. Lisäksi käyrän tulisi olla aidosti konvekssi, jotta tiheysfunktio pysyy positiivisena koko toteutushintajoukossa (Breedon ja Litzenberger 1978, 627). Kaikkina aineistoon kuuluvina päivinä eivät molemmat ehdot täyty, jolloin tiheysfunktioista tulee virheellinen.

Aineiston diskreettiydestä johtuen toisen derivaatan approksimaatiota lähde-tään hakemaan diskreetillä toisen kertaluvun differenssillä muodostamalla kaa-van (3.24) mukainen portfolio. Differenssioperaation seurauksena menetetään muo-dostettavasta tiheysfunktioista sekä toteutushintavektorin ensimmäinen että viimei-nen havainto. Näin ollen tiheysfunktioiksi saadaan

$$f(\mathbf{K})|_{K=S_T} = \mathbf{w}\mathbf{C} \Leftrightarrow \begin{bmatrix} f(K_2) \\ f(K_3) \\ \vdots \\ f(K_{M-1}) \end{bmatrix}_{(M-2) \times 1} = \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 & -2 & 1 \end{bmatrix}_{(M-2) \times M} \begin{bmatrix} C(K_1, T) \\ C(K_2, T) \\ \vdots \\ C(K_M, T) \end{bmatrix}_{M \times 1}$$

missä  $f(\cdot)$  on implisiittinen riskineutraali tiheysfunktio,  $\mathbf{w}$  on portfolion tridiago-naalinen painomatriisi,  $K_m$  on  $m$ :s toteutushinta järjestyksessä pienimmästä suu-rimpaan ja  $M$  on toteutushintojen määrä. Matriisien alakulmiin on merkitty niiden dimensiot selkeyttämään laskuoperaatiota.

Tiheysfunktio on määritetty edellä maturiteetissa, joten se on vielä diskontat-tava nykyhetkeen. Hetken  $t$  tiheysfunktio on siis lopulta muotoa

$$e^{-r_f \Delta t} f(\mathbf{K}).$$

## 5 TULOKSET JA JOHTOPÄÄTÖKSET

Tässä tutkielmassa keskitytään kolmen eri aikahorisontin arviointiin. Ensiksi käsitellään lyhyttä aikaväliä, joka kattaa jäljellä olevan juoksuaajan kahdesta viikosta kuukauteen, toisena käsitellään keskipitkää aikajännettä, joka käsittää kolmesta neljään kuukautta ennen option erääntymistä ja lopuksi vielä pitkällä tähtäimellä, kun juoksu-aikaa on jäljellä 10–14 kuukautta. Estimoitu implisiittinen jakauma on havainnollistettavuuden vuoksi piirretty jokaiselta päivältä. Jakaumia arvioidaan visuaalisen tarkastelun perusteella, mutta tarkemmassa analyysissä on hyödynnetty jakaumien momenteja siltä osin kuin aineistossa on havaintopisteitä. Vertailukohtana pidetään tutkimuskysymyksen 1 ja käsitellyn teoreettisen viitekehysten mukaisesti Black–Scholes–Merton-mallin implikoimaa normaalijakaumaa. Malliyhtälöstä (3.18) voidaan analyttisesti määrittää kaavan (3.25) mukainen toisen kertaluvun derivaatta toteutushinnan suhteen, joka vastaa tässä yhteydessä log-normaalijakauman tiheysfunktioita

$$f(K) = \frac{\partial^2 C(t, K)}{\partial K^2} = e^{-r_f(T-t)} \frac{\phi(L(t, K))}{K \sigma \sqrt{T-t}}$$

missä  $\phi(\cdot)$  on standardin normaalijakauman tiheysfunktio.

Momentit on määritetty integraaleina

$$\begin{aligned} m_1 &= \mu = \mathbb{E}[S_T] = \int_{-\infty}^{\infty} [K \cdot f(K)] dK \Big|_{K=S_T} \\ m_2 &= \sigma^2 = \mathbb{E}[(S_T - \mu)^2] = \int_{-\infty}^{\infty} [(K - \mu)^2 \cdot f(K)] dK \Big|_{K=S_T} \\ m_i &= \mathbb{E}\left[\left(\frac{S_T - \mu}{\sigma}\right)^i\right] = \int_{-\infty}^{\infty} \left[\left(\frac{K - \mu}{\sigma}\right)^i \cdot f(K)\right] dK \Big|_{K=S_T}, \quad i = 3, 4, 5, 6 \end{aligned}$$

joista  $m_3$  on jakauman vinous (engl. *skewness*) ja  $m_4$  huipukkuus (engl. *kurtosis*). Korkeammat momentit  $m_5$  ja  $m_6$  painottavat jakauman häntiä enemmän kuin normaalit vinous- ja huipukkuustunnusluvut. Käytettävässä tiheysfunktiossa otetaan huomioon aineiston rajallisuus eli asetetaan  $f(K) = 0$ , kun  $K \notin [\min(K), \max(K)]$ , eli toisin sanoen jakauman häntien  $K \in [0, \min(K))$  ja  $K \in (\max(K), \infty)$  realisoitumistodennäköisyydet asetetaan eroavuuksien tarkastelussa nolnaan. Näin ollen integraali määritetään kunakin päivänä ja kullakin erääntymispäivällä käytännössä vain välillä  $[\min(K), \max(K)]$ , koska rajallinen aineisto ei mahdollista jakauman häntien estimointia. Tällä tavalla määritetyt momentit eivät siis vastaa tavanomaisia jakauman momenteja.

Koska saatavilla olevat  $\min(K)$  ja  $\max(K)$  vaihtelevat eri päivinä ja erääntymispäivinä, ei määritettyjä momenteja pystytä suoraan vertailemaan eri ajankohtina. Sen sijaan tutkimuskysymyksen 1 kohdalla voidaan tutkia momenttien erotusta jakaumien välillä. Nollahypoteesina oletetaan, että jakaumien momenttien välillä ei

ole eroa  $H_0 : \Delta m_i = 0, i = 1, \dots, 6$ , ja hypoteeseja testataan yksinkertaisesti yhden otoksen  $t$ -testillä. Tutkimuskysymyksen 2 osalta jakaumien eroavuuden stabiiliutta ajan suhteen tutkitaan visuaalisin havainnoin jakaumista ja niiden momenttien erotuksista.

Jakaumien eroavuuden mittana käytetään lisäksi integroitua keskineliövirhettä (engl. *root mean integrated squared error*, RMISE), kuten esimerkiksi muiden muassa Bu ja Hadri (2007), Lee (2014), Santos ja Guerra (2015) sekä Kundu, Kumar ja Tomar (2019). Tämä virhemitta vastaa 2-normia ja määritellään seuraavasti:

$$\text{RMISE} = \sqrt{\mathbb{E} \left[ \int_{-\infty}^{\infty} [\hat{f}(S_T) - f(S_T)]^2 dS_T \right]},$$

missä  $\hat{f}(\cdot)$  on estimoidun implisiittisen jakauman tiheysfunktio ja  $f(\cdot)$  on normaalijakauman mukainen tiheysfunktio (Scott 2015, 41). Integroituja keskineliövirheitä eri aikaväleillä verrataan  $t$ -testeillä.

Poikkeuksellisesti muutamana päivänä, erityisesti koronaviruksen puhjettua ja aiheutettua korkeaa volatiliteettia, menetelmä epäonnistui pahasti ja jakauman muoto oli varsin erikoinen (ks. liite B). Markkinatilanteesta johtuen optioiden hinnoissa ja sitä kautta myös implisiittisessä volatiliteetissa esiintyi suurta hajontaa, kuten esimerkikuviossa 8b. Suuri hajonta vaikeutti implisiittisen volatiliteetin tasoitusalgoritmin toimintaa niin paljon, että se ei kyennyt tasoittamaan käyrää riittävästi, vaan käyrän sovitteseen jäi liikaa kohinaa. Näin ollen implisiittisen volatiliteettisovituksen perusteella palautettu hintafunktio oli muodoltaan epämääräinen, minkä takia myös sen toisesta derivaatasta muodostui erikoinen tiheysfunktio.

**Taulukko 4:** Estimoitujen jakaumien momenttien eroavuus normaalijakaumasta**(a)** Lyhyt aikaväli

Momentti	Estimaatti	Keskivirhe	$t$ -testisuure	$p$ -arvo	
$\Delta\mu$	44,439	7,469	5,950	< 0,001	***
$\Delta\sigma$	-10,883	5,662	-1,922	0,058	
$\Delta m_3$	-2,307	0,106	-21,838	< 0,001	***
$\Delta m_4$	9,333	0,834	11,187	< 0,001	***
$\Delta m_5$	-83,740	8,686	-9,640	< 0,001	***
$\Delta m_6$	719,729	98,867	7,280	< 0,001	***

**(b)** Keskipitkä aikaväli

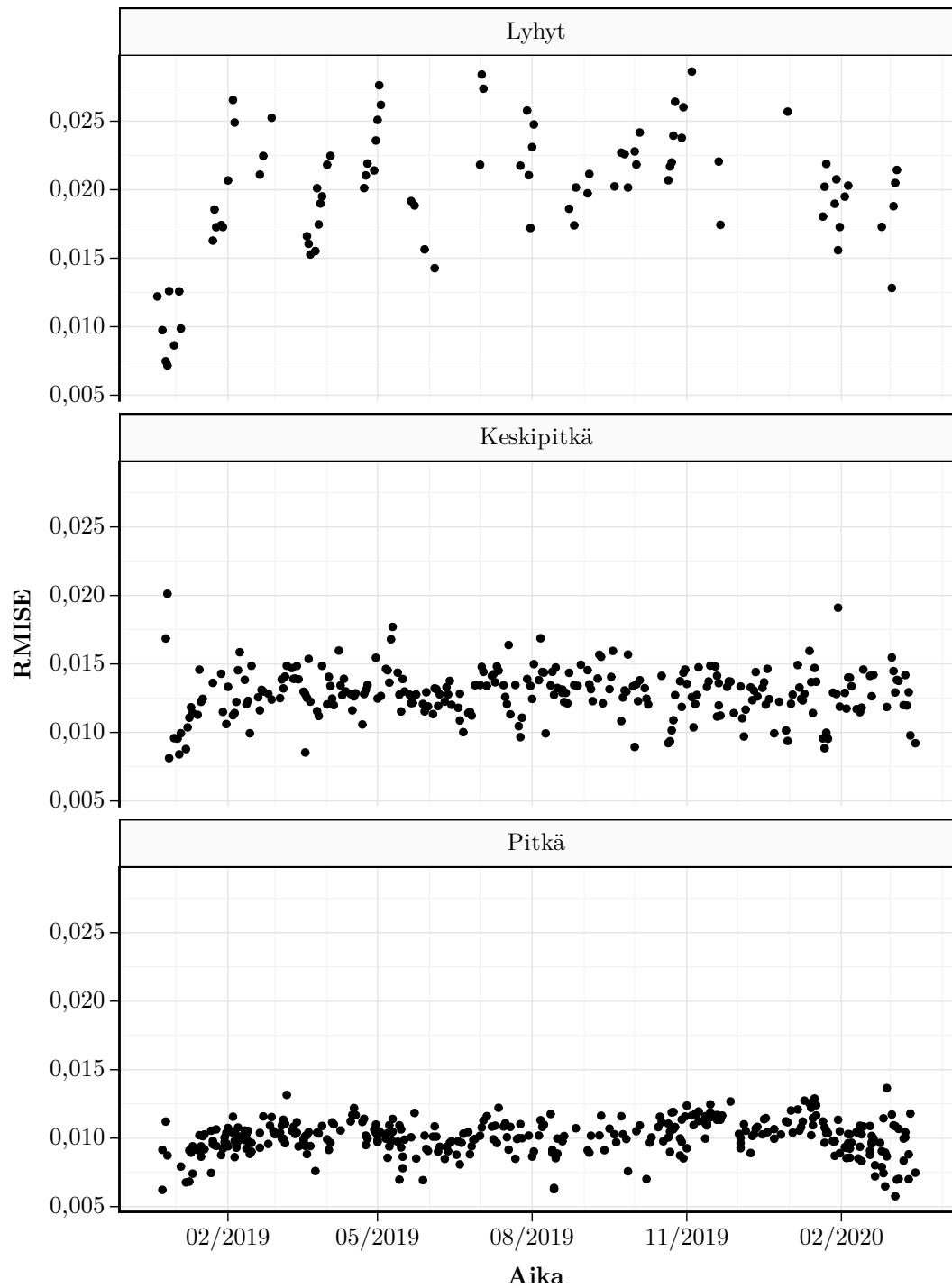
Momentti	Estimaatti	Keskivirhe	$t$ -testisuure	$p$ -arvo	
$\Delta\mu$	4,312	9,652	0,447	0,655	
$\Delta\sigma$	36,097	3,978	9,074	<0,001	***
$\Delta m_3$	-0,705	0,047	-15,127	<0,001	***
$\Delta m_4$	0,809	0,093	8,687	<0,001	***
$\Delta m_5$	-7,340	0,725	-10,122	<0,001	***
$\Delta m_6$	25,541	4,109	6,216	<0,001	***

**(c)** Pitkä aikaväli

Momentti	Estimaatti	Keskivirhe	$t$ -testisuure	$p$ -arvo	
$\Delta\mu$	-128,053	12,418	-10,312	<0,001	***
$\Delta\sigma$	139,454	5,449	25,593	<0,001	***
$\Delta m_3$	-0,242	0,021	-11,570	<0,001	***
$\Delta m_4$	-0,159	0,027	-5,922	<0,001	***
$\Delta m_5$	-2,079	0,182	-11,425	<0,001	***
$\Delta m_6$	-0,016	0,424	-0,038	0,970	

Käytetyt merkitsevyystasot:

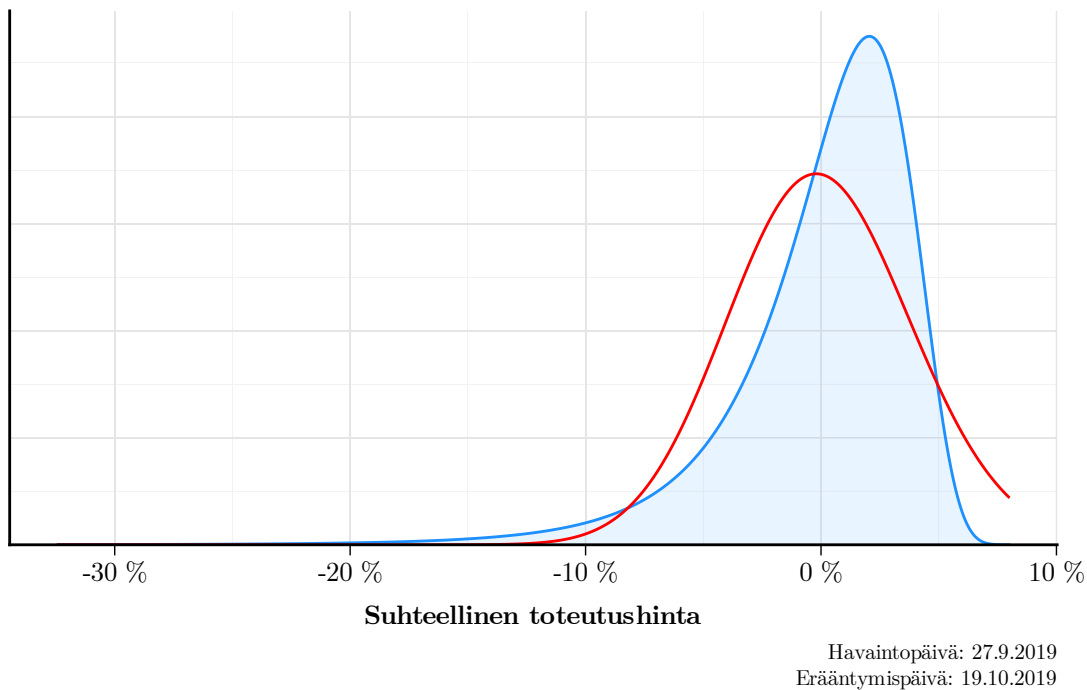
\*\*\*:  $p$ -arvo alittaa merkitsevyystason 0,001



**Kuvio 10:** Jakaumien eroavuuksien integroidut keskineliövirheet eri aikahorisonteilla

## 5.1 Lyhyt aikaväli

Lyhyen aikavälin tarkastelussa käytetty tutkimusmenetelmä onnistui estimoimaan implisiittisen jakauman suurimmassa osassa havaintopäiviä hyvin. Hyvin onnistunutta implisiittistä jakaumaa on havainnollistettu yhtenä päivänä kuviossa 11. Tyypillisesti jakauma on visuaalisesti tarkasteltuna sekä huipukkaampi kuin lognormaalijakauma että negatiivisesti vino, kuten esimerkkikuviossa. Taulukosta 4a voidaan havaita myös momenttien eroavuuksien perusteella, että kaikissa tutkituissa kuudessa ensimmäisessä momentissa on hyvin selvää tilastollisesti merkitsevää eroa, paitsi keskihajonnassa. Momentit  $m_3$ – $m_6$  vahvistavat visuaaliset havainnot negatiivisesta vinoudesta ja korkeammasta huipukkaudesta. Vastaus tutkimuskysymykseen 1 on lyhyen aikavälin osalta, että jakauma poikkeaa selvästi oletusarvoisesta normaalijakaumasta.



**Kuvio 11:** Lyhyen juoksuaajan optioiden implikoima todennäköisyysjakauma

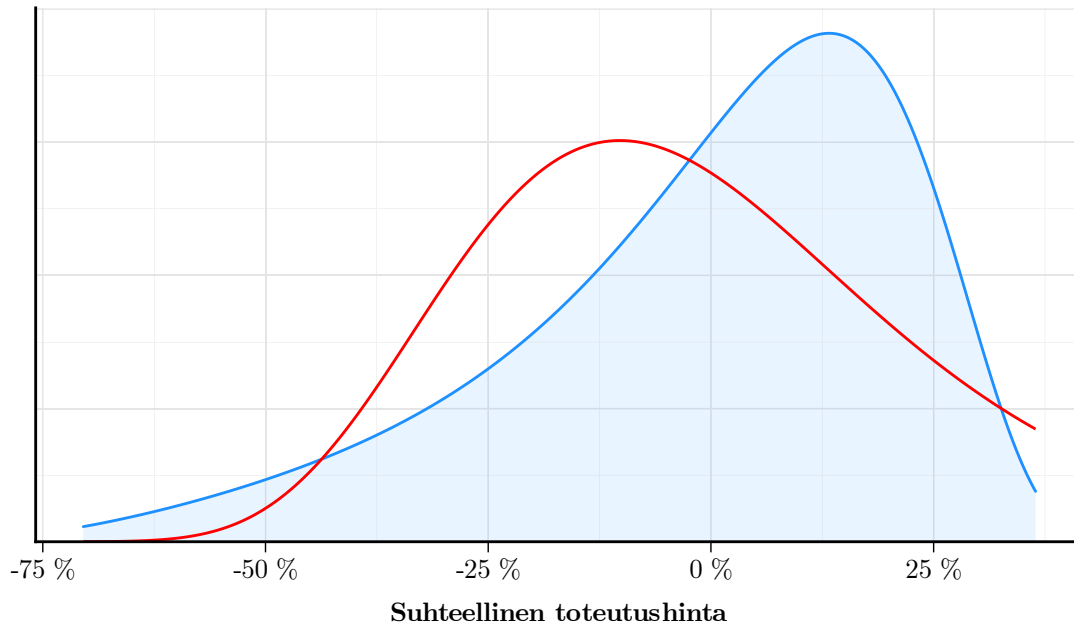
Jakauman pysyvyyttä tarkastellessa voidaan todeta, että sen poikkeavuus lognormaalijakaumasta on jopa yllättävän samanlaista, vaikka tarkasteluajakaorizontti ulottuu yli vuoden ajalle. Momenttien erotuksia (ks. liitteen C kuvio (a)) tarkastellessa havaitaan, ettei minkään momentin erotuksessa ole trendiä tai selkeää kausivaihtelua. Momenttien erotuksista löytyy kuitenkin joitakin selkeitä poikkeamia, mutta ne ovat pääasiassa lyhytaikaisia ja palaavat takaisin keskimääräiselle tasolle varsinkin korkeammassa momentissa. Kuitenkin keskiarvossa ja keskihajonnassa esiintyy pari suurempaa poikkeamaa kesän 2019 ja vuodenvaihteen 2019–2020

aikana. Odotusarvojen ero kasvaa kummassakin tapauksessa ja pysyy korkeana yhdestä kahteen kuukautta. Samalla ero keskihajonnassa muuttuu negatiiviseksi. Erot normalisoituvat kuitenkin takaisin keskimääräiselle tasolle. Kuviossa 10 esitetyn integroidun keskineliövirheen perusteella eroavuuden määrä on pysynyt alun muutama havaintoa lukuun ottamatta keskimäärin samalla tasolla. Tutkimuskysymyksen 2 osalta vastaus on, että suhteessa normaalijakaumaan estimoitu implisiittinen jakauma pysyy muodoltaan kohtuullisen vakiona ajan suhteen.

## 5.2 Keskipitkä aikaväli

Keskipitkän aikavälin kohdalla estimointimenetelmä epäonnistui enää hyvin harvoissa tapauksissa. Visuaalisesti tarkasteltuna jakaumalla on edelleen keskimäärin pienempi vinous ja suurempi huipukkuus kuin normaalijakaumalla. Taulukon 4b perusteella jakaumien implikoimat odotusarvot eivät aineiston asettamat rajoitteet huomioiden eroa tilastollisesti merkitsevästi toisistaan. Sen sijaan muissa momenteissa eroavuutta löytyy. Tarkasteluaikavälillä implisiittinen jakauma näyttäisi, että markkinat odottavat korkeampaa volatilitteettia kuin normaalijakauma olettaa. Korkeammat momentit osoittavat tyypilliseen tapaan, että markkinoilla jakaumassa olisi havaittavissa suurempaa negatiivista vinoutta ja paksumpia häntiä kuin normaalijakaumassa. Kerroinestimaatit eivät poikke erityisen rajusti nolasta, mutta pienten keskivirheiden takia erot ovat tilastollisesti hyvin merkitseviä. Tutkimuskysymyksen 1 kohdalla keskipitkä aikaväli osoittaa lyhyen aikavälin tapaan, että estimoitu implisiittinen jakauma poikkeaa normaalijakaumasta.

Jakauman muoto suhteessa normaalijakaumaan vaikuttaa momenttien erotusten (ks. liitteen C kuvio (b)) perusteella pysyvän keskipitkällä aikavälillä hyvin vakiona. Yhdessäkään tarkasteltavassa momentissa ei esiinny trendiä tai selkeää kausivaihtelua. Vaihtelu näyttää hyvin satunnaiselta. Korkeimmissa momenteissa näyttäisi olevan joitakin korkeampia piikkejä, mutta korkeat momentit ovat muutenkin herkkiä äärihavainnoille, joten poikkeukselliset havainnot voi laskea satunnaisuuden aiheuttamiksi. Momenttien  $m_3$ – $m_6$  erotuksissa voidaan havaita maaliskuussa 2019 hetkittäinen poistuma keskimääräiseltä tasolta. Tällöin jakaumassa on ollut lievästi enemmän vinoutta vasemmalle sekä korkeampaa huipukkuutta, tosin ero on vain pientä. Integroitu keskineliövirhe pysyy kuvion 10 perusteella keskimäärin samalla tasolla eli eroavuus on suuruusluokaltaan samanlaista koko aikajaksolla. Tämän takia tutkimuskysymyksen 2 vastaus keskipitkän aikavälin osalta on, että jakaumassa ei ole tapahtunut ajan mittaan juurikaan poikkeuksia normaalijakaumaan nähden, vaan ero on pysynyt lähes vakiona.



Havaintopäivä: 13.3.2020  
Eräätymispäivä: 20.6.2020

**Kuvio 12:** Keskipitkän juoksuajan optioiden implikoima todennäköisyysjakauma

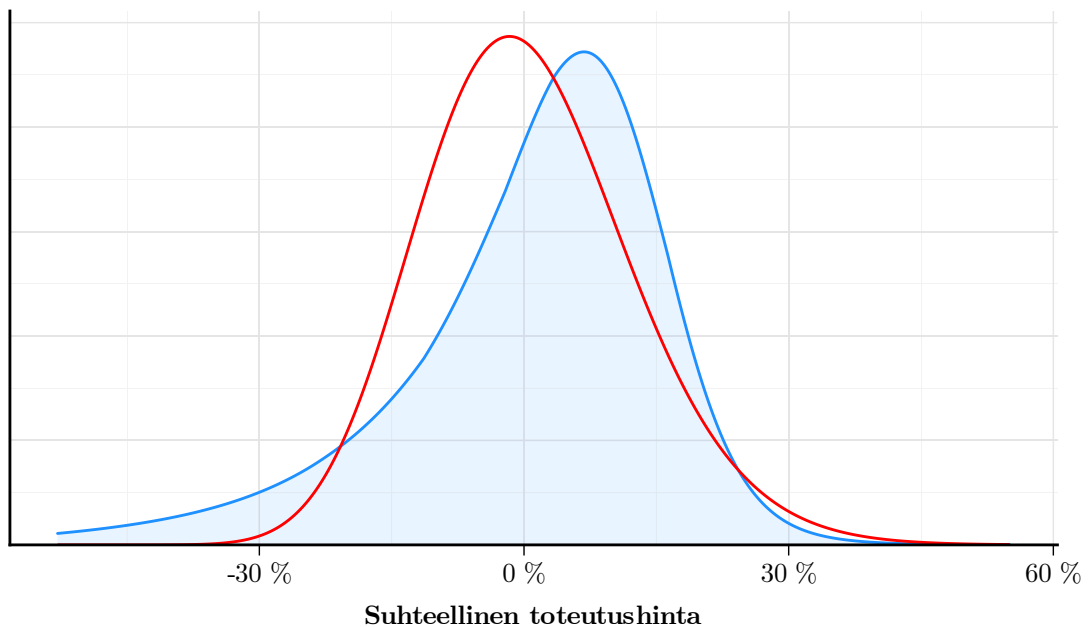
Kuviossa 12 on koronaviruksen puhkeamisen jälkeen yhden päivän esimerkkijakauma. Siitä voidaan havaita, että vasemman hännän äärimmäisen negatiiviset realisoitumistodennäköisyydet ylittävät normaalijakauman erittäin paljon, kun taas positiivisella puolella häntä on selvästi ohuempi, vaikkakin se katkeaa harmittavan aikaisin. Samalla ero odotusarvoissa on kasvanut, koska implisiittisellä jakaumalla se on noin 360 indeksipistettä korkeampi kuin normaalijakaumalla.

### 5.3 Pitkä aikaväli

Pisimmällä noin vuoden juoksuajalla estimointi onnistui jälleen lähes kaikkina tarkastelupäivinä, mutta aineiston tarjoama väli  $[\min(K), \max(K)]$  on nyt selvästi kapeampi kuin muilla aikahorisonteilla. Visuaalisesti tarkasteltuna estimoitu implisiittinen jakauma alkaa lähestyä normaalijakaumaa. Taulukossa 4c esitettyjen momenttien erotusten perusteella keskiarvo ja keskihajonta poikkeavat kovasti jakaumien välillä, mutta korkeammat momentit ovat suuruusluokaltaan melko lähellä toisiaan. Kuudetta momenttia lukuun ottamatta erotukset ovat kuitenkin tilastollisesti merkitseviä, vaikka kerroinestimaatit ovatkin itseisarvoltaan melko pieniä. On merkittävää huomata, että pitkällä aikavälillä ero huipukkuuksien välillä on kääntynyt negatiiviseksi eli implisiittisellä jakaumalla on tunnuslukujen valossa epätyypillisesti ohuempat hännät kuin normaalijakaumalla. Tutkimuskysymyksen 1 osalta erot jakaumien välillä eivät ole enää niin suuria pitkällä aikavälillä, mutta implisiitti-

nen jakauma eroaa silti tilastollisesti merkitsevästi normaalijakaumasta erityisesti kahden ensimmäisen momentin osalta.

Normaalijakaumaan verrattuna estimoidun jakauman muoto pysyy momenttien erotusten (ks. liitteen C kuvio (c)) perusteella jälleen lähes vakiona, koska momenttien erotuksista ei löydy trendiä tai selkää kausivaihtelua. Pitkällä aikavälillä momenttien erotuksissa ei ole havaittavissa edes yksittäisiä suuria poikkeamia, mutta momenteissa  $m_3$ – $m_6$  vaihtelu on heteroskedastista. Vuoden 2019 maaliskuu–kesäkuussa ja syys–lokakuussa vaihtelu momenttien erotuksissa on selvästi pienempää kuin alkuvuonna 2019 ja 2020. Erotusten odotusarvo pysyy kuitenkin näinäkin aikoina vakiona. Alkuvuoden 2020 tilannetta voi selittää koronaviruksen aiheuttama turbulenssi, kun markkinoilla odotukset ovat vaihdelleet rajusti. Laskumarkkinan aikana pitkän aikavälin implisiittinen odotusarvo on kuitenkin kääntynyt pienoiseen nousuun normaalijakaumaan verrattuna. Tämä kertoo markkinoiden luottamuksesta, että indeksi palautuu pidemmällä välillä takaisin nousuun. Samalla keskihajonta on kääntynyt pienemmäksi kuin normaalijakaumalla. Tuleva kehitys jakauman muodosta jää nähtäväksi, mutta tällä aineistolla tutkimuskysymyksen 2 vastaus on, että keskimäärin ero normaalijakaumaan on pysynyt vakiona, vaikka päivittäistä hajontaa on ilmenyt jonkin verran. Integroitu keskineliövirhe (kuvio 10) tukee tätä tulkintaa, koska poikkeavuus on pysynyt samassa kokoluokassa koko tarkasteluajavälillä.



Havaintopäivä: 8.2.2019  
Erääntymispäivä: 19.12.2019

**Kuvio 13:** Pitkän juoksuajan optioiden implikoima todennäköisyysjakauma

Kuviossa 13 on erään päivän estimoitu jakauma, jossa havaintopisteitä on runsaammin kuin keskimäärin pitkällä aikavälillä. Selvästi paksumpi vasen häntä implisiittisessä jakaumassa aiheuttaa sen, että myös sen odotusarvo on matalampi kuin normaalijakaumalla, vaikka jakaumien moodit ovatkin päinvastoin.

#### 5.4 Erot eri aikavälien kesken

Tutkimuskysymyksen 3 kannalta on tärkeää käsitellä, millaisia eroja estimoiduilla implisiittisillä jakaumilla on havaittavissa. Jakauman poikkeamista normaalijakaumasta mitataan tässä tutkielmassa integroidulla keskineliövirheellä. Eri aikahorisonteilla poikkeavuudet normaalijakaumasta olivat erilaisia. [Welchin \(1947\)](#) kahden otoksen  $t$ -testillä testataan, ovatko integroidut keskineliövirheet keskimäärin yhtä suuria kaikissa kolmessa otoksessa. Keskiarvojen yhtäsuuruuden testauksessa käytetään [Welchin](#) testiä, koska ei voida tehdä vakuuttavaa oletusta eri otosten yhtä suurista variansseista.

**Taulukko 5:** Keskimääräinen integroitu keskineliövirhe eri aikahorisonteilla

$t$ -testisuure	<b>Keskipitkä</b>	<b>Pitkä</b>
$p$ -arvo	$\mu = 0,0129$	$\mu = 0,0101$
<b>Lyhyt</b>	-13,873	-19,655
$\mu = 0,0199$	< 0,001 ***	< 0,001 ***
<b>Keskipitkä</b>		-22,060
$\mu = 0,0129$		< 0,001 ***

Käytetyt merkitsevyystasot:

\*\*\*:  $p$ -arvo alittaa merkitsevyystason 0,001

Jo kuvion 10 perusteella on selvää, että lyhyen aikavälin integroitu keskineliövirhe on keskimäärin suurempaa kuin muiden tarkasteluhorisonttien kohdalla. Tämä tulos voidaan vahvistaa taulukon 5 ensimmäisen rivin perusteella. Vaikka lyhyen aikavälin keskiarvo poikkeaaakin absoluuttisesti eniten muista aikahorisonteista, lyhyen aikavälin integroidussa keskineliövirheessä on kuitenkin niin paljon hajontaa, että tilastollisen merkitsevyyden näkökulmasta  $t$ -testisuureet ja siten myös  $p$ -arvot ovat eroavuuden luotettavuuden kannalta heikompia kuin keskipitkän ja pitkän aikavälin välillä. Tilastollinen merkitsevyys on siitä huolimatta hyvin selvää. Keskipitkän ja pitkän aikavälin kesken absoluuttinen ero integroitujen keskineliövirheiden keskiarvoissa on pienempi, mutta huomattavasti pienemmän hajonnan vuoksi ero on tilastollisesti selvästi merkitsevin. Tuloksista siis selviää, että aikahorisontilla on vaikutusta implisiittisen jakauman eroavuuteen oletuksena käytetystä normaalijakaumasta.

Erot eri aikahorisonttien välillä voidaan ainakin osittain selittää momenttien eroavuuksien perusteella. Taulukosta 4 nähdään, että lyhyellä aikavälillä odotusarvo on korkeampi kuin normaalijakaumalla, mutta pitkällä aikavälillä taas selvästi pienempi. Keskipitkällä aikavälillä tilastollisesti merkitsevää eroa ei syntynyt. Keskihajonnan kohdalla muutokset ovat päinvastaisia: lyhyellä aikavälillä eroa normaalijakaumaan ei ole tilastollisesti merkitsevällä tavalla, mutta pitkällä aikavälillä keskihajonta on implisiittisellä jakaumalla suurempaa kuin normaalijakauma olettaa. Viinouden osalta ero normaalijakaumaan pienenee aikavälin kasvaessa. Huipukkuuden osalta tapahtuu merkittävä muutos, koska lyhyellä ja keskipitkällä aikavälillä implisiittinen jakauma osoittaa paksuhäntäisyyttä, mutta pitkällä aikavälillä momentin ero kääntyykin negatiiviseksi eli normaalijakauma onkin implisiittisetä jakaumaa paksuhäntäisempi, mikä ei ole tyypillistä. Korkeampien momenttien absoluuttiset erot pienenevät aikahorisontin kasvaessa.

## 5.5 Tulosten merkitys optioiden hinnoittelun kannalta

Black–Scholes–Merton-hinnoittelukaavat (3.18) ja (3.19) ovat varmasti rahoituksessa eniten käytettyjen hinnoittelumallien joukossa. Mutta kuten luvussa 3.4 käsiteltiin, mallissa on melko rajoittavat oletukset, jotka vaikuttavat sen luotettavuuteen oikeilla markkinoilla. Tulokset osoittavat, että niiden optioiden kohdalla, joiden toteutushinta on lähellä nykyistä indeksin tasoa, hinnoitteluvirhe mallia käytettäessä on melko pientä, mutta mitä kauemmas toteutushinta siirtyy kohti häntiä, sitä enemmän hinnoitteluvirhettä esiintyy.

Hinnoitteluvirhettä tutkitaan molemmilla optiotyypeillä, kaikilla aikaväleillä sekä kolmessa eri kategoriassa: lähellä toteutushintaa olevat optiot (engl. *at-the-money option*, lyhennetään tässä ATM), perusarvoa sisältävät optiot (engl. *in-the-money option*, lyhennetään tässä ITM) ja perusarvon kannalta arvottomat optiot (engl. *out-of-the-money option*, lyhennetään tässä OTM). Perusarvolla tarkoitetaan yhtälöiden (3.1) ja (3.2) kuvaamaa palkkiota, jos optio erääntyisi välittömästi tarkasteluhetkellä. ATM-optioksi lasketaan tässä yhteydessä sellaiset optiot, joiden toteutushinta poikkeaa enintään 2 % indeksin tasosta.

Hinnoitteluvirheeseen sovelletaan prosentuaalista keskivirhettä (engl. *mean percentage error*, MPE)

$$\begin{aligned} \text{MPE}_C &= \mathbb{E} \left[ \frac{C_{\text{BSM}} - C}{C} \right] \\ \text{MPE}_P &= \mathbb{E} \left[ \frac{P_{\text{BSM}} - P}{P} \right] \end{aligned}$$

joka määritellään tässä Black–Scholes–Merton-mallin mukaisen hinnan  $C_{\text{BSM}}$  ( $P_{\text{BSM}}$ ) ja markkinahinnan  $C$  ( $P$ ) erotuksen perusteella. Kahdella havainnolla markkinahinta on niin lähellä nollaa, että prosentuaalinen virhe kasvaa erittäin suureksi (yli

200 000 %). Erittäin poikkeavat havainnot puolestaan nostavat keskivirhettä, minkä takia nollahypoteesit jäävät voimaan kahdessa luokassa vain näiden kahden havainnon takia. Tästä syystä nämä kaksi havaintoa on jätetty arvioinnin ulkopuolelle.

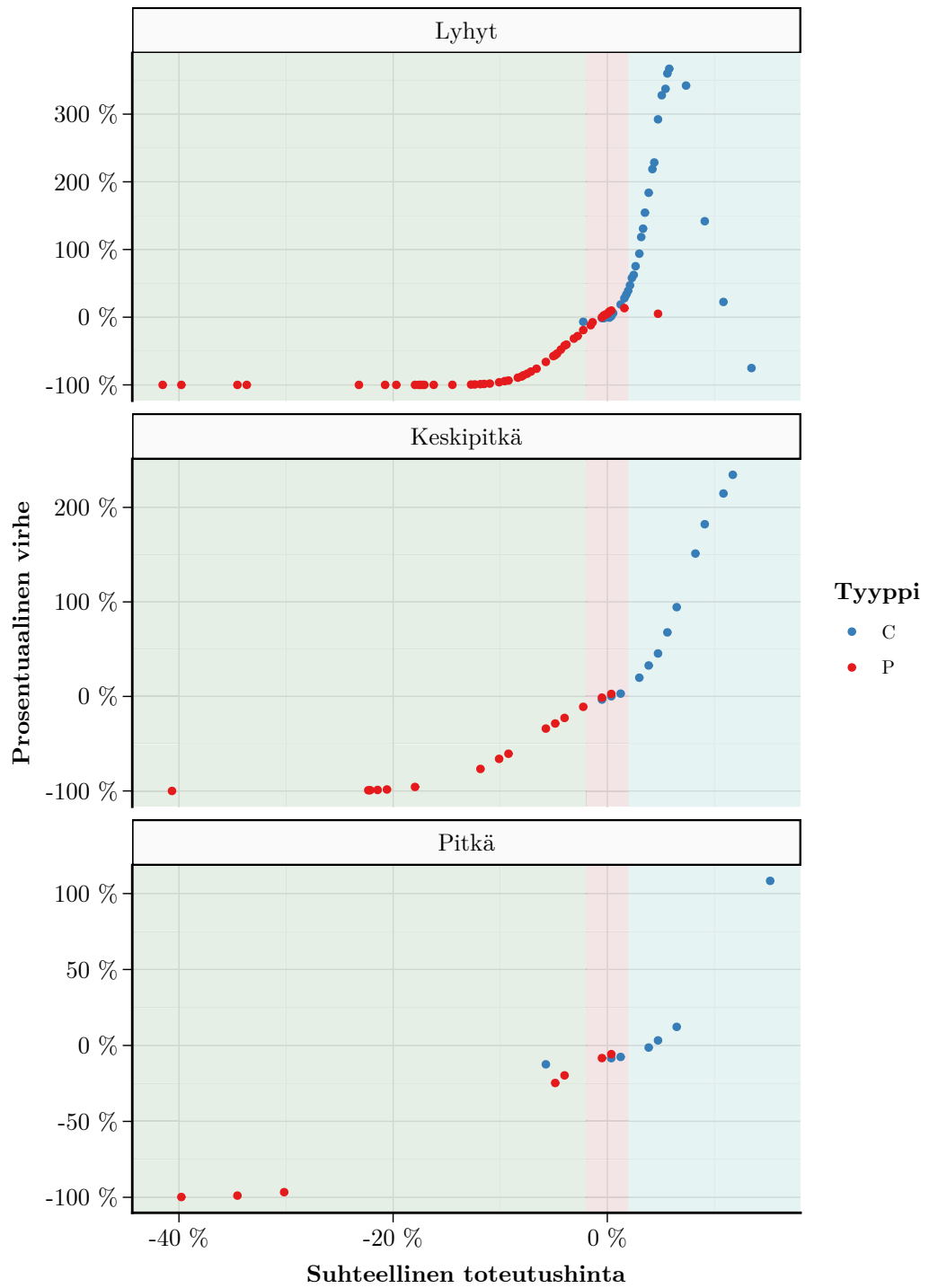
**Taulukko 6:** Optiohintojen prosentuaalinen keskivirhe eri kategorioissa

Kategoria	Horisontti	Tyyppi	MPE, %	<i>t</i> -testisuure	<i>p</i> -arvo	
ITM	Lyhyt	Osto	-2,424	-4,533	< 0,001	***
		Myynti	0,441	0,672	0,502	
	Keskipitkä	Osto	-4,241	-5,169	< 0,001	***
		Myynti	11,827	13,808	< 0,001	***
	Pitkä	Osto	-11,732	-11,630	< 0,001	***
		Myynti	7,311	4,601	< 0,001	***
ATM	Lyhyt	Osto	28,910	24,196	< 0,001	***
		Myynti	3,030	5,930	< 0,001	***
	Keskipitkä	Osto	-2,659	-7,647	< 0,001	***
		Myynti	-5,520	-16,137	< 0,001	***
	Pitkä	Osto	-13,485	-22,863	< 0,001	***
		Myynti	-11,558	-19,469	< 0,001	***
OTM	Lyhyt	Osto	229,339	35,724	< 0,001	***
		Myynti	-81,635	-194,981	< 0,001	***
	Keskipitkä	Osto	158,768	17,592	< 0,001	***
		Myynti	-75,413	-181,828	< 0,001	***
	Pitkä	Osto	102,333	8,967	< 0,001	***
		Myynti	-67,693	-110,595	< 0,001	***

Käytetyt merkitsevyystasot:

\*\*\*: *p*-arvo alittaa merkitsevyystason 0,001

Optioiden hintojen erotukset 18 kategoriassa on esitetty taulukossa 6. Kuten koko tämän tutkielman empiirisessä osiossa, myös tässä yhteydessä tulkinta heijastelee sitä ajatusta, että optioiden markkinahinnat ovat virheettömiä ja hinnoittelun virhe syntyy yksinomaan Black–Scholes–Merton-mallin liian pelkistävästä oletuksista. Tutkimuskysymyksen 4 kannalta on merkittävää huomata, että hinnoitteluvirhe on yhtä kategorialuokkaa lukuun ottamatta tilastollisesti hyvin merkitsevää ja monissa tapauksissa taloudellisesti ajateltuna vielä hyvinkin merkittävää, koska poikkeamat ovat useissa kategorioissa valtavia. Seuraavaksi käsittelemme hinnoitteluvirheiden kooka ja suuntaa eri kategorioissa, mistä kuviossa 14 on esimerkkinä tuloksia hyvin kuvaava tilanne.



**Kuvio 14:** Prosentuaalinen hintavirhe optioiden hinnoissa

Suurimmat hinnoitteluvirheet tapahtuvat OTM-optioiden joukossa sekä ostetta myyntioptioilla. Osto-optioiden hinta Black–Scholes–Merton-mallin mukaan voi lyhyellä aikavälillä olla keskimäärin jopa yli kolminkertainen markkinahintaan verrattuna. Pidemmillä aikajaksoilla virhe ei ole aivan yhtä suuri, mutta silti pitkälläkin aikavälillä hinnan tulisi olla mallin mukaan noin kaksinkertainen. Syy tähän löytyy juuri siitä, että markkinoilla vallitseva implisiittinen todennäköisyysjakauma on taulukon 4 mukaan vasemmalle vino. Paksuhäntäisyys korostaa vaikutusta lyhyellä ja keskipitkällä aikavälillä. Tällöin todennäköisyys sille, että kohde-etuusindeksi nousee ja sitä kautta OTM-osto-optiolla on arvoa erääntymispäivänä, jää pieneksi. Black–Scholes–Merton-mallin pohjalla oleva normaalijakauma antaa suuremman painon nousulle, minkä vuoksi myös mallin mukainen option hinta heijastelee parempaa nousupotentiaalia. Tämänsuuntainen hinnoitteluvirhe on optioita ostavan sijoittajan kannalta vaarallinen, koska suosittu ja markkinoilla laajasti käytetty hinnoittelumalli näyttää antavan viitteitä siitä, että optiot ovat markkinoilla reilusti alihinnoiteltuja, vaikka todellisuudessa alhainen hinta heijastelee huonoja tuottomahdollisuuksia.

Myyntioptioiden kohdalla tulokset ovat täysin päinvastaisia, mutta hinnoitteluvirheet ovat suuruusluokaltaan yhä hyvin merkittäviä. Lyhyellä aikavälillä hintaero on tässäkin tapauksessa suurempi kuin pitkällä aikavälillä. Syy, miksi OTM-myyntioptiot ovat markkinoilla kalliimpia kuin Black–Scholes–Merton-mallin mukaan voisi olettaa, on siinä, että nyt kohde-etuusindeksin laskulla on implisiittisen jakauman mukaan suurempi todennäköisyys kuin normaalijakaumalla johtuen juuri negatiivisesta vinoudesta. Kuvion 14 perusteella Black–Scholes–Merton-mallin mukaan OTM-myyntioptioiden tulisi olla käytännössä arvottomia, kun toteutushinta poikkeaa indeksin nykytasosta riittävän paljon alaspäin. Hintavirhe laskee selkeästi kohti -100 %:n tasoa lyhyellä aikavälillä toteutushinnan erotessa noin 15 % indeksin tasosta alaspäin. Kaupankäyntiaktiivisuutta tarkasteltaessa sekä myynti- että ostooptioissa ATM-optiot ovat suosituimpia lyhyellä ja keskipitkällä aikavälillä. Lyhyen ja pitkän aikavälin OTM-myyntioptiot ovat selkeästi suositumpia kuin ITM-optiot, mikä voi kertoa siitä, että spekulatiiviset sijoittajat yrittävät hyödyntää edellä kuvattua hintavirhettä. OTM-optioiden käyttö spekulatiivisessa sijoittamisessa on mielekästä siitä syystä, että niillä ei ole yhtälöiden (3.1) ja (3.2) mukaista perusarvoa, vaan niiden arvo perustuu puhtaasti aika-arvoon.

ITM-optioilla hinnoitteluvirheet eivät ole enää yhtä räikeitä kuin OTM-optioilla. Ostooptioilla Black–Scholes–Merton-mallin mukaiset hinnat ovat keskimäärin hieinan matalampia kuin markkinahinnat, mutta virhe kasvaa aikavälin pidentyessä. Pitkällä aikavälillä hinnoitteluvirhe on kuitenkin vain noin -11,7 %. Myyntioptioiden puolella hinnoitteluvirhe on vastakkaiseen suuntaan ja muuten samaa suuruus-

luokkaa osto-optioiden kanssa, mutta saatujen tulosten valossa lyhyellä aikavälillä ei esiinny tilastollisesti merkitsevää virhettä.

ATM-optioiden osalta tulokset ovat moninaisia. Alkuperäisen oletuksen mukaan hinnoitteluvirheen tulisi olla pienintä ATM-optioiden kohdalla, mutta näin käy ainoastaan keskipitkällä aikavälillä. Lyhyellä aikavälillä myyntioptioiden virhe on vielä pientä, mutta osto-optioiden kohdalla virhe on yllättävän korkea, lähes 30 %. Virheen suunta ei myöskään ole yksikäsitteinen: lyhyellä aikavälillä Black–Scholes–Merton-mallin hinnat ovat liian korkeita, mutta keskipitkällä ja pitkällä aikavälillä puolestaan liian alhaisia.

## 6 YHTEENVETO

Tämän tutkielman tarkoituksena on ollut selvittää, millainen markkinoilla vallitseva implisiittinen todennäköisyysjakauma on, miten se käyttäytyy ajan suhteen ja millaisia eroavuuksia siinä ilmenee. Jakauman muodon merkitystä on korostettu optioiden hinnoittelun näkökulmasta. Tavoitteisiin on lähdetty vastaamaan tutkimalla ensin matemaattisesti, miten satunnaisuus on aiemmassa tutkimuksessa nähty ja millaisia stokastisia prosesseja osakkeet noudattavat. Tämän perusteella on käsitelty ehdollisten vaateiden hinnoittelun periaatteet. Edettäessä yksinkertaisista yhden periodin binomimarkkinoista jatkuva-aikaiseen normaalijakaumaan on optiohinnoittelussa hyödynnetty matemaattisen rahoituksen tärkeimpiä työkaluja, mukaan lukien osakkeen hintaprosessia kuvaava geometrinen Brownin liike, Cameron–Martin–Girsanovin mittamuunnosteoria ja Itô–Döblinin lause. Näillä välineillä on pyritty tarkentamaan optioiden hinnoittelua vaihe kerrallaan. Lopulta on esitetty optiomarkkinat mullistaneet Black–Scholes–Merton-hinnoittelukaavat osto- ja myyntioptioille.

Vaikka teoreettisesta näkökulmasta kaikki edellä esitetyt tulokset ovat hyvin perusteltavissa, on empiirisissä tutkimuksissa havaittu poikkeavuuksia. Esimerkiksi perinteisen teorian riippumattomuusoletus hinnanmuutoksissa on osoittautunut puutteelliseksi. Fraktionaalissa Brownin liikkeessä esiintyvän Hurstin parametrin on estimoitu olevan yli 0,5, mikä tarkoittaa, että hintaprosessilla on pitkä muisti. Tämä yhdessä muiden melko rajoittavien oletusten kanssa johtaa siihen, että markkinoilla suositussa käytössä olevat Black–Scholes–Merton-hinnoittelukaavat eivät toimi empiirisesti.

Mainituissa Black–Scholes–Merton-kaavoissa volatiliteettia lukuun ottamatta kaikki muut hinnoitteluun vaikuttavat parametrit ovat saatavilla suoraan markkinoilta. Koska volatiliteetti on ainut vapaa parametri, ottamalla myös option hinta markkinoilta voidaan määrittää implisiittinen volatiliteetti. Useinkaan implisiittinen volatiliteetti ei ole vakio eri toteutushinnoilla, vaan volatiliteetti on toteutushinnan suhteen konvekssi, jolloin puhutaan volatiliteettihymystä. Jos volatiliteetti on lisäksi myös laskeva, voidaan puhua volatiliteettivinousesta. Tällaiset poikkeamat implikoivat, että Black–Scholes–Merton-kaavat eivät päde. Useimmiten virheen suuruus korostuu niillä optioilla, joiden toteutushinta on kaukana kohde-etuuden nykyisestä hinnasta. Virhe todistaa myös sen, että normaalijakauma ei ole tuoton todellinen jakauma oikeilla markkinoilla.

Optioiden hintojen tarjoamalla informaatiolla markkinoilla vallitseva implisiittinen todennäköisyysjakauma voidaan määrittää optioiden hintafunktion toisesta derivaatasta toteutushinnan suhteen. Tätä estimointitehtävää on lähestytty approksimoimalla luonnollinen kuutiosplini volatiliteettihymykuvaajaan. Approksimaatio-

funktion avulla on palautettu takaisin optioiden hintafunktio, josta toinen derivaatta on määritetty. Implisiittinen tiheysfunktio on estimoitu jokaiselle päivälle ja kaikille aineistossa tarjolla oleville erääntymispäiville.

Tulokset osoittavat, että lyhyellä aikavälillä implisiittisen jakauman eroavuus normaalijakaumaan on suurinta. Eroavuutta on tutkittu integroidun keskineliövirheen ja momenttien eroavuuden avulla. Integroitu keskineliövirhe on keskimäärin 0,0199, ja momentit osoittavat, että markkinoilla yleisesti todetut negatiivinen vinous ja paksuhäntäisyys ovat voimassa. Jakaumien eroavuus integroidun keskineliövirheen suhteen on vähäisempää keskipitkällä (RMISE 0,0129) ja pitkällä (RMISE 0,0101) aikavälillä. Myös momenttien perusteella eroavuus laimenee aikavälin kasvaessa ja itse asiassa huipukkuuden osalta erotus kääntyy hieman negatiiviseksi. Tulokset ovat tilastollisesti selvästi merkitseviä muutamaa yksittäistä momenttia lukuun ottamatta.

Saavutettujen tulosten merkitys korostuu Black–Scholes–Merton-mallin aiheuttamia hinnoitteluvirheitä tutkittaessa. Perusarvoltaan arvottomat OTM-optiot sisältävät suurimmat virheet: osto-optioilla Black–Scholes–Merton-mallin hinnat ovat aikavälistä riippuen kaksi tai jopa yli kolme kertaa markkinahintoja korkeampia. Tällainen virhe yhdessä yleisimmin käytetyssä hinnoittelumallissa voi olla äärimmäisen vakava optioita ostavan sijoittajan kannalta, jos tämä luottaa malliin liiaksi. Myyntioptioilla virhe on täysin päinvastainen, mutta suuruusluokaltaan edelleen merkittävä. Virheet voidaan perustella implisiittisen jakauman muodolla: negatiivinen vinous painottaa kohde-etuuden laskun todennäköisyyksiä, jolloin osto-option arvo markkinoilla on normaalijakaumaan pohjautuvaa Black–Scholes–Merton-mallia matalampi ja myyntioption puolestaan taas korkeampi.

Perusarvoa sisältävillä ITM-optioilla hinnoitteluvirheet ovat melko maltillisia verrattuna OTM-optioihin. Tuloksista selviää, että ITM-osto-optioiden kohdalla virhe kasvaa juoksuajan pidentyessä. Lisäksi lyhyen aikavälin ITM-myyntioptioiden kohdalla Black–Scholes–Merton-mallia hyödynnettäessä tilastollisesti merkitsevää hinnoitteluvirhettä ei synny. Option toteutushinnan ollessa lähellä indeksin nykytasoa ei tuloksissa ole nähtävissä selkeää linjaa. Näillä ATM-optioilla hinnoitteluvirheen tulisi olla pienintä, mutta tämä toteutuu vain keskipitkällä aikavälillä.

Tämän tutkielman tuloksista selviää siis, että sijoittaja voi altistua merkittävillekin hinnoitteluvirheille hyödyntäessään Black–Scholes–Merton-mallia optioiden hinnoittelussa. Estimoitujen implisiittisten jakaumien muodot vastaavat yleistä käsitystä negatiivisesta vinoudesta ja paksuhäntäisyydestä. Vaikka käytetyssä aineistossa esiintyy paikoittain huomattavaakin hajontaa, ovat tulokset tilastollisen merkitsevyyden kannalta selviä lähes kaikilla mittareilla. Vain joitain yksittäisiä eroavuuksia ei kyetty osoittamaan.

Tulosten yleistettävyyttä tukee se, että tutkielmassa käytetty aineisto pohjautuu Yhdysvaltojen suurimpien yhtiöiden muodostamaan indeksiin, johon kuuluu 500 yhtiötä. Käytettyyn S&P 500 -indeksiin liittyvä aineisto on myös laadukkaampaa kuin esimerkiksi pienen Suomen tai edes Euroopan pääindekseihin liittyvät indeksioptiot, koska ensinnäkin optioita on tarjolla enemmän ja toiseksi niiden kaupankäyntiaktiivisuus on huomattavasti parempaa. Tässä törmätään kuitenkin myös siihen ongelmaan, että vastaavanlainen tutkimus saattaa olla vaikea toteuttaa tai epäluotettava sellaisilla markkinoilla, joilla ei yhtä hyvää aineistoa ole saatavilla. Validiteetin suhteen yhden tietyn markkinan käyttäminen aineistossa rajoittaa tulosten pätevyyttä, koska eri markkinoilla on erilaisia piirteitä. Myös melko lyhyt aineiston aikajänne ei välttämättä edusta markkinoiden kaikkia ominaisuuksia, jotka tulisivat esiin pidemmällä tähtäimellä. Aikavälin valinnassa tulee kuitenkin ottaa huomioon, että tutkimusmenetelmä on laskennallisesti melko vaativa ja päivänsisäiseen käyttöön soveltumaton.

Reliabiliteetin kannalta aiemmista tutkimuksista on poimittu parhaimpia käytäntöjä esimerkiksi valitsemalla ei-parametrinen tutkimusmenetelmä ja tarkentamalla approksimointia luonnollisella kuutiosplinillä paraabelin sijaan. Toisaalta valittua menetelmää ei ole arvioitu tai vertailtu jonkin rinnakkaisen menetelmän kanssa, mikä johtuu tutkielman laajuuden aiheuttamasta rajoitteesta. Lisäksi kuten alan tutkimuksessa yleisestikin, markkinoilla vallitsevat epätäydellisyydet, mukaan lukien kaupankäyntikustannukset ja verotus, voivat aiheuttaa oman häiriönsä. Seuraavaksi voisikin olla tarpeellista vertailla menetelmien eroja, tutkia jakauman muotoa pidemmällä aikavälillä ja yrittää mallintaa markkinoiden epätäydellisyyksiä.

## LÄHTEET

- Arrow, K. J. – Debreu, G. (1954) Existence of an equilibrium for a competitive economy. *Econometrica*, Vol. 22 (3), 265–290.
- Bachelier, L. (1900) The theory of speculation. *Annales Scientifiques de l'École Normale Supérieure*, Vol. 3 (17), 21–86. (Alkuteos Théorie de la spéculation, käännös May, D. 2011).
- Bender, C. – Sottinen, T. – Valkeila, E. (2011) Fractional processes as models in stochastic finance. Teoksessa: *Advanced Mathematical Methods for Finance*, toim. Di Nunno, G. – Øksendal, B. Springer, Berlin, Heidelberg, 75–103.
- Björk, T. (2009) *Arbitrage theory in continuous time*. 3. p. Oxford University Press, New York.
- Black, F. – Scholes, M. (1972) The valuation of option contracts and a test of market efficiency. *The Journal of Finance*, Vol. 27 (2), 399–417.
- Black, F. – Scholes, M. (1973) The pricing of options and corporate liabilities. *Journal of Political Economy*, Vol. 81 (3), 637–654.
- Bliss, R. R. – Panigirtzoglou, N. (2002) Testing the stability of implied probability density functions. *Journal of Banking and Finance*, Vol. 26 (2–3), 381–422.
- Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, Vol. 31 (3), 307–327.
- Borowski, K. (2018) Testing 65 equity indexes for normal distribution of returns. *Journal of Economics and Management*, Vol. 34 (4), 5–38.
- Breeden, D. T. – Litzenberger, R. H. (1978) Prices of state-contingent claims implicit in option prices. *The Journal of Business*, Vol. 51 (4), 621–651.
- Bru, B. – Yor, M. (2002) Comments on the life and mathematical legacy of Wolfgang Doeblin. *Finance and Stochastics*, Vol. 6 (1), 3–47.
- Bu, R. – Hadri, K. (2007) Estimating option implied risk-neutral densities using spline and hypergeometric functions. *The Econometrics Journal*, Vol. 10 (2), 216–244.
- Campa, J. M. – Chang, P. H. K. – Reider, R. L. (1998) Implied exchange rate distributions: evidence from OTC option markets. *Journal of International Money and Finance*, Vol. 17 (1), 117–160.

- Capinski, M. – Zastawniak, T. (2003) *Mathematics for Finance: An Introduction to Financial Engineering*. Springer, Berlin, Heidelberg.
- Corrado, C. J. – Su, T. (1996) Skewness and kurtosis in S&P 500 index returns implied by option prices. *Journal of Financial Research*, Vol. 19 (2), 175–192.
- Cox, J. C. – Ross, S. A. – Rubinstein, M. (1979) Option pricing: A simplified approach. *Journal of Financial Economics*, Vol. 7 (3), 229–263.
- Damodaran, A. (2020) S&P earnings: 1960–current. <[http://pages.stern.nyu.edu/~adamodar/New\\_Home\\_Page/datafile/spearn.htm](http://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/spearn.htm)>, haettu 30.3.2020.
- Engle, R. F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, Vol. 50 (4), 987–1007.
- Fama, E. F. (1965) The behavior of stock-market prices. *The Journal of Business*, Vol. 38 (1), 34–105.
- Fama, E. F. (1970) Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, Vol. 25 (2), 383–417.
- Federal Reserve Bank of St. Louis (2020) FRED. <<https://fred.stlouisfed.org>>, haettu 17.3.2020.
- Godfrey, M. D. – Granger, C. W. J. – Morgenstern, O. (1964) The random-walk hypothesis of stock market behavior. *Kyklos*, Vol. 17 (1), 1–30.
- Gomes, L. M. P. – Soares, V. J. S. – Gama, S. M. A. – Matos, J. A. O. (2018) Long-term memory in Euronext stock indexes returns: An econophysics approach. *Business and Economic Horizons*, Vol. 14 (4), 862–881.
- Grith, M. – Härdle, W. K. – Schienle, M. (2012) Nonparametric estimation of risk-neutral densities. Teoksessa: *Handbook of Computational Finance*, toim. Duan, J.-C. – Härdle, W. K. – Gentle, J. E. Springer, Berlin, Heidelberg, 277–305.
- Harezlak, J. – Ruppert, D. – Wand, M. P. (2018) *Semiparametric Regression with R*. Springer, New York.
- Heston, S. L. (1993) A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, Vol. 6 (2), 327–343.

- Hull, J. – White, A. (1987) The pricing of options on assets with stochastic volatilities. *The Journal of Finance*, Vol. 42 (2), 281–300.
- Hull, J. C. (2011) *Options, Futures, and Other Derivatives*. 8. p. Pearson.
- Investopedia (2020) What is the average annual return for the S&P 500? <<https://www.investopedia.com/ask/answers/042415/what-average-annual-return-sp-500.asp>>, haettu 13.5.2020.
- Itô, K. (1944) Stochastic integral. *Proceedings of the Imperial Academy*, Vol. 20 (8), 519–524.
- Itô, K. (1946) On a stochastic integral equation. *Proceedings of the Japan Academy*, Vol. 22 (2), 32–35.
- Jackwerth, J. C. – Rubinstein, M. (1996) Recovering probability distributions from option prices. *The Journal of Finance*, Vol. 51 (5), 1611–1631.
- Jamdee, S. – Los, C. A. (2007) Long memory options: LM evidence and simulations. *Research in International Business and Finance*, Vol. 21 (2), 260–280.
- Jegadeesh, N. – Titman, S. (1993) Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, Vol. 48 (1), 65–91.
- Keim, D. B. (1985) Dividend yields and stock returns: Implications of abnormal January returns. *Journal of Financial Economics*, Vol. 14 (3), 473–489.
- Kundu, A. – Kumar, S. – Tomar, N. K. (2019) Option implied risk-neutral density estimation: A robust and flexible method. *Computational Economics*, Vol. 54 (2), 705–728.
- Kwok, Y.-K. (2008) *Mathematical Models of Financial Derivatives*. 2. p. Springer, Berlin, Heidelberg.
- Lee, S. H. (2014) Estimation of risk-neutral measures using quartic B-spline cumulative distribution functions with power tails. *Quantitative Finance*, Vol. 14 (10), 1857–1879.
- Liu, Z. – Moghaddam, M. D. – Serota, R. A. (2019) Distributions of historic market data – stock returns. *The European Physical Journal B*, Vol. 92 (3), 60.
- Malkiel, B. G. (2003) The efficient market hypothesis and its critics. *Journal of Economic Perspectives*, Vol. 17 (1), 59–82.

- Malkiel, B. G. (2019) *A Random Walk Down Wall Street*. 12. p. W. W. Norton & Company, New York.
- Mandelbrot, B. B. (1963) The variation of certain speculative prices. *The Journal of Business*, Vol. 36 (4), 394–419.
- Mandelbrot, B. B. – Van Ness, J. W. (1968) Fractional Brownian motions, fractional noises and applications. *SIAM Review*, Vol. 10 (4), 422–437.
- Merton, R. C. (1973) Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, Vol. 4 (1), 141–183.
- Mills, T. C. (1995) Modelling skewness and kurtosis in the London stock exchange FT-SE index return distributions. *Journal of the Royal Statistical Society. Series D (The Statistician)*, Vol. 44 (3), 323–332.
- Monteiro, A. M. – Tütüncü, R. H. – Vicente, L. N. (2008) Recovering risk-neutral probability density functions from options prices using cubic splines and ensuring nonnegativity. *European Journal of Operational Research*, Vol. 187 (2), 525–542.
- Ncube, M. – Satchell, S. (1997) The statistical properties of the Black–Scholes option price. *Mathematical Finance*, Vol. 7 (3), 287–305.
- R Core Team (2020) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <<https://www.r-project.org>>.
- Ross, S. A. (1976) Options and efficiency. *The Quarterly Journal of Economics*, Vol. 90 (1), 75–89.
- Ross, S. M. (2011) *An Elementary Introduction to Mathematical Finance*. 3. p. Cambridge University Press.
- Salazar Celis, O. – Liang, L. – Lemmens, D. – Tempère, J. – Cuyt, A. (2015) Determining and benchmarking risk neutral distributions implied from option prices. *Applied Mathematics and Computation*, Vol. 258, 372–387.
- Santos, A. – Guerra, J. (2015) Implied risk neutral densities from option prices: Hypergeometric, spline, lognormal, and Edgeworth functions. *Journal of Futures Markets*, Vol. 35 (7), 655–678.
- Scott, D. W. (2015) *Multivariate Density Estimation: Theory, Practice, and Visualization*. 2. p. John Wiley & Sons, Hoboken, New Jersey.
- Shimko, D. (1993) Bounds of probability. *Risk*, Vol. 6 (4), 33–37.

- Taboga, M. (2016) Option-implied probability distributions: How reliable? How jagged? *International Review of Economics & Finance*, Vol. 45, 453–469.
- Tsay, R. S. – Chen, R. (2019) *Nonlinear Time Series Analysis*. John Wiley & Sons, Hoboken, New Jersey.
- Welch, B. L. (1947) The generalization of Student's problem when several different population variances are involved. *Biometrika*, Vol. 34 (1–2), 28–35.
- Wickham, H. (2010) A layered grammar of graphics. *Journal of Computational and Graphical Statistics*, Vol. 19 (1), 3–28.
- Yan, H. – Han, L. (2019) Empirical distributions of stock returns: Mixed normal or kernel density? *Physica A: Statistical Mechanics and its Applications*, Vol. 514, 473–486.

## LIITTEET

### A Analyysissä käytetyt ohjelmistot ja R-paketit

#### (a) Pilvipalvelin

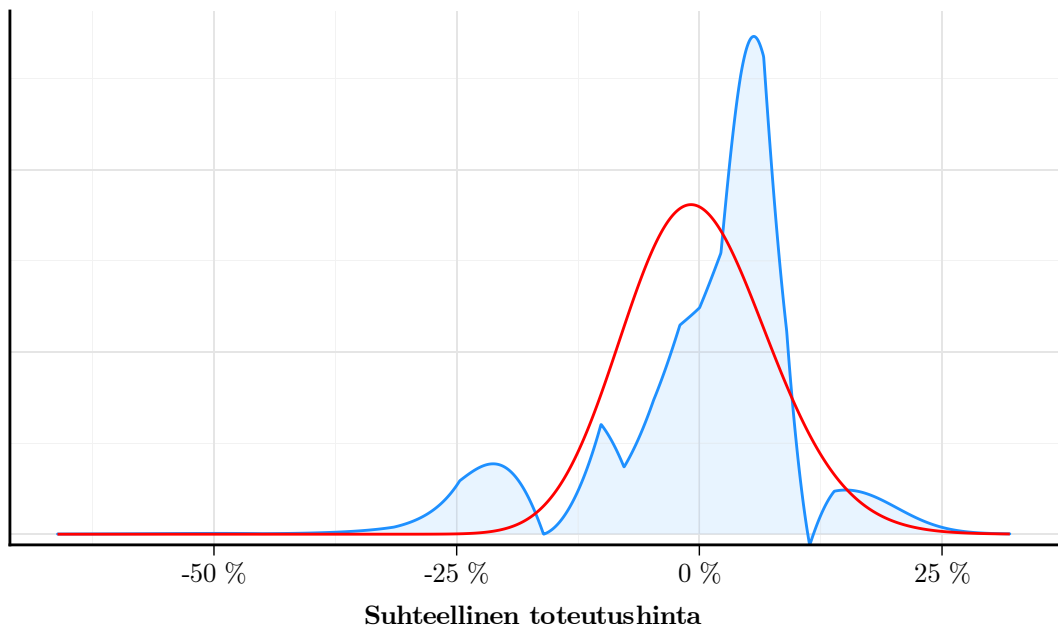
Käyttöjärjestelmä	Ubuntu Server 20.04
R	4.0.1
RStudio Server	1.3.959
Apache	Välityspalvelimena paikallisesti ajetulle RStudio Serverille

#### (b) R-paketit

Paketin nimi	Käyttötarkoitus
<code>tidyverse</code>	Sisältää monia hyödyllisiä paketteja aineiston käsittelyyn
– <code>readr</code>	Aineiston nopea tuonti muista tiedostoista
– <code>tibble</code>	Uusi taulukko-objektimuoto, jossa <code>tidyverse</code> -pakettien avulla datan muokkaus on helppoa
– <code>dplyr</code>	Aineiston helppo manipulointi
– <code>purrr</code>	Riveittäin tapahtuvien operaatioiden ajo
– <code>stringr</code>	Tekstimuotoisten muuttujien manipulointi
– <code>ggplot2</code>	Datan visualisointi <i>Grammar of Graphics</i> -periaatetta noudattaen (Wickham 2010)
<code>magrittr</code>	Komentojen ketjutusoperaattorit
<code>lubridate</code>	Aikamuuttujien laskenta
<code>scales</code>	<code>ggplot2</code> -kuvaajien akselien muokkaus
<code>Cairo</code>	Kuvaajien tulostaminen PDF-tiedostoiksi
<code>extrafont</code>	$\LaTeX$ -fontti kuvaajiin
<code>data.table</code>	Riveittäin tehtävien operaatioiden tehokkaampi suoritus
<code>foreach</code> ja <code>doParallel</code>	R:n ajaminen prosessorin yhden säikeen sijaan kaikilla ytimillä
<code>xtable</code>	R:n tulosteiden vienti $\LaTeX$ -muotoon
<code>somebm</code>	Fraktionaalisen Brownin liikkeen simulointi
<code>broom</code>	Tilastollisten testien tallentamien objektien muokkaus <code>tidyverse</code> -yhteesopiviksi

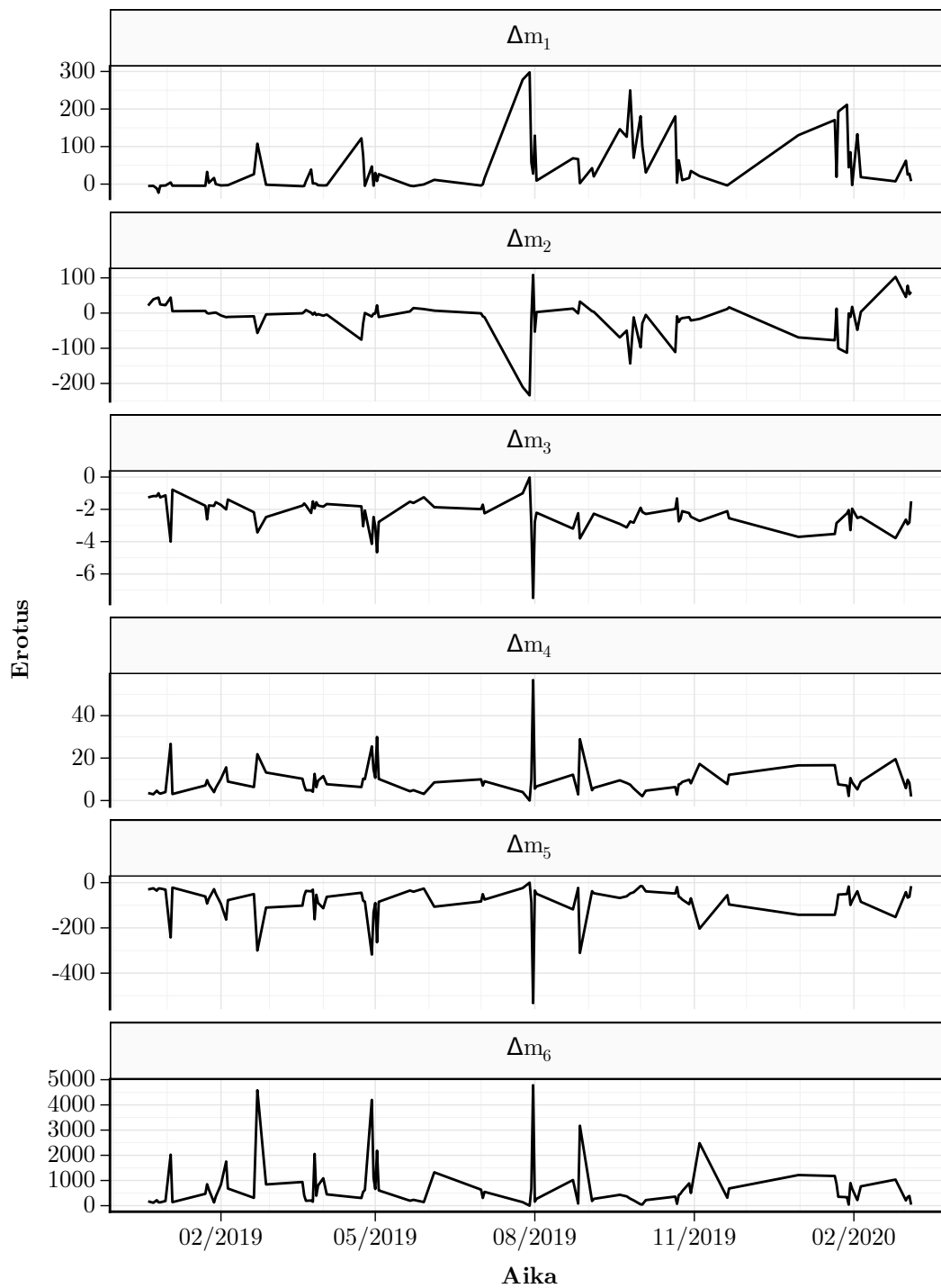
## B Epäonnistunut jakauman estimointi

Implisiittisen jakauman estimointi ei kaikkina havaintopäivinä onnistunut kovin hyvin. Esimerkkinä tästä lyhyen aikavälin jakauma 28.2.2020, jolloin jakaumassa esiintyi useita epänormaaleja huippuja ja negatiivisia todennäköisyyksiä. Molemmat merkit viittaavat siihen, että aineistossa on ollut paljon hajontaa ja implisiittisen volatiliiteetin tasoitus on ollut vaikeaa algoritmille. Tästä on aiheutunut optioiden hintafunktioon erikoisia hyppyjä, minkä vuoksi sen derivaatat ovat muodoltaan sekavia.

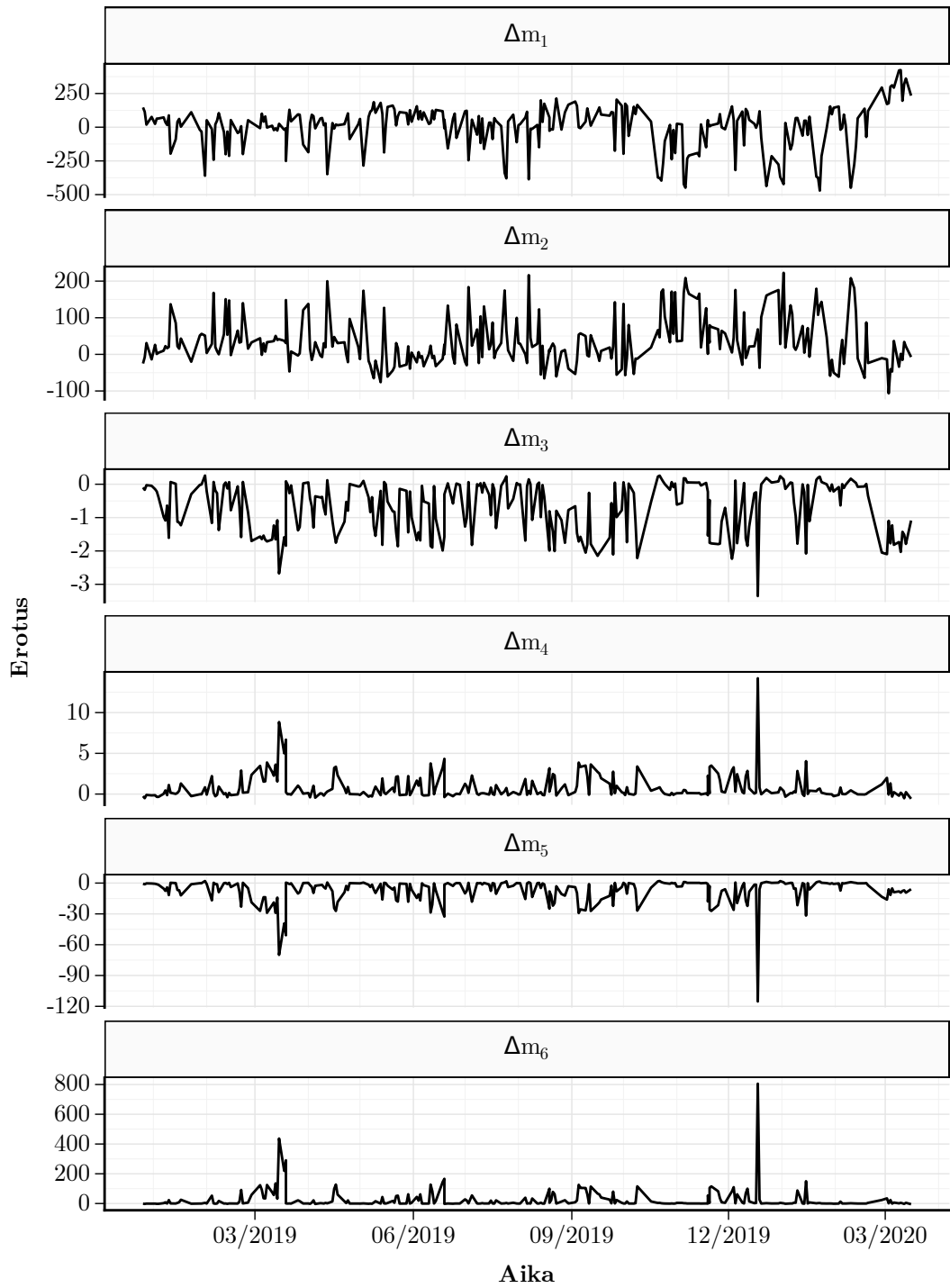


Havaintopäivä: 28.2.2020  
Erääntymispäivä: 20.3.2020

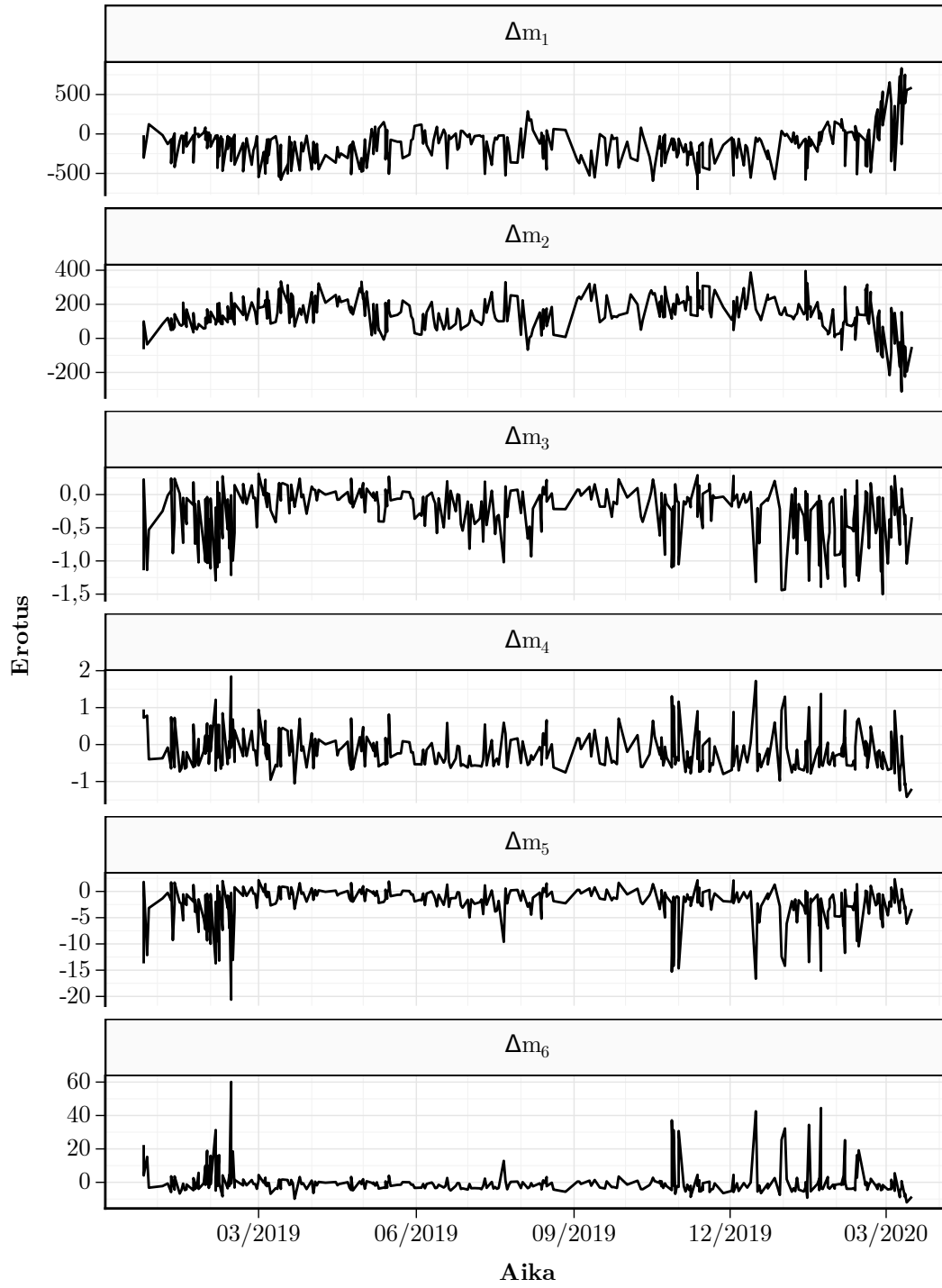
## C Estimoitujen jakaumien momenttien ero normaalijakaumaan



(a) Lyhyt aikaväli



(b) Keskipitkä aikaväli



(c) Pitkä aikaväli