

Aalto-yliopisto

Teknillisen fysiikan ja matematiikan koulutusohjelma

Systeemianalyysin laboratorio

22.1.2015

Ville Venojärvi

LOTTO RAHAPELINÄ SEKÄ SEN MYYNNIN MAL- LINTAMINEN JA SIMULOINTI

Työn valvoja: Prof. Harri Ehtamo

Työn ohjaaja: Di. Harri Järvinen

Työn saa tallentaa ja julkaista Aalto-yliopiston avoimilla verkkosivuilla. Muilta osin kaikki oikeudet pidätetään.

The purpose of this thesis has primarily been to investigate factors affecting gaming of Lotto by mathematical models and to create a simulation model which can be used to study the expected yearly sales of Lotto and its variance. With the simulation model it was also studied what kind of Jackpot accumulation cycles would be feasible for Lotto. The law aspect, mathematical structure and the history of Finnish Lotto and lotto games in general are also briefly discussed. There is also a short review at the studies made about mathematical modelling of lotto games.

Mathematical modelling was conducted by forming linear multiple regression models for different durations of Lotto (1, 2, 3, 5, 10 weeks). The regression models were done by regular linear multiple regression models as well as logistic regression models. Modelling was done with two different approaches to increase the validity of the results.

Lotto gaming was found to depend on many factors, all of which could not be included in the regression models (eg. marketing investments and weather). Most important factor proved to be the jackpot of Lotto and its size. The effect of jackpot proved to be greater with shorter durations and it was perceived to clearly be non-linear for one week games. Other significant predictors were for example season of the year and the count how many times the jackpot has rolled over. Mathematical models for the sales were most successful for one week durations, in general it can be said that the shorter the duration the better the mathematical model performed.

With the simulation models it was discovered that on a yearly basis Lotto sales varies tens of millions of euros (magnitude 10% of yearly sales) depending on which kind of jackpots happen to occur in a given calendar year. This was one of the most important results of this thesis. It was also discovered that the conditional jackpot offerings greatly affect the yearly sales.

The results achieved by mathematical models and simulations can be as a rule, considered to be reliable. When interpreting the results it has to be remembered that all the factors affecting the gaming of Lotto could not be included in the models and some of the factors were discovered after the completion of this study. Although the factors affecting gaming of Lotto can be expected to change only slowly in time, it might be interesting to study how the results would change if the data of new Lotto game would be added to analysis.

Tämän diplomityön tarkoituksena on ollut ensisijaisesti tutkia Loton pelaamiseen vaikuttavia tekijöitä matemaattisten mallinnosten avulla sekä muodostaa simulaattorimalli, jonka avulla on pyritty tutkimaan Loton odotusarvoisia vuositason myyntejä sekä niiden vaihteluita. Simulaattorimallin avulla pyrittiin myös tutkimaan, millaisin edellytyksin Lotossa voitaisiin erilaisia potteja tarjota (pottitarjoama) pelaajille. Edellä mainittujen lisäksi työssä on tehty lyhyt katsaus niin Loton lakitekniiseen perustaan, rahapelirakenteeseen kuin Loton sekä lottopelien historiaan yleisesti. Työssä on myös lyhyt katsaus julkaisuihin, joissa on käsitelty lottopelien pelaamisen mallintamista matemaattisin menetelmin.

Pelaamisen mallintamista tutkittiin muodostamalla usean selittäjän regressiomallit Loton eri kestojen (1, 2, 3, 5, 10 viikkoa) myynneille. Regressiomallit toteutettiin sekä tavanomaisina usean selittäjän regressiomalleina, että logistiseen regressioon pohjautuvina malleina. Kahdella eri lähestymistavalla haluttiin pääasiassa parantaa tulosten luotettavuutta.

Loton pelaamisen havaittiin riippuvan erittäin monesta tekijästä, joita kaikkia ei voitu regressiomalleihin sisällyttää (esim. markkinointipanostukset ja säätila). Pelaamisen merkittävimmäksi selittäjäksi havaittiin Loton potti ja sen suuruus. Loton potin vaikutus pelaamisen on sitä voimakkaampaa, mitä lyhyemmästä kestoista on kyse ja sen vaikutuksen huomattiin olevan selkeästi epälineaarista kertapelien (yhden viikon kestot) osalta. Muita merkittäviä selittäjiä pelaamiselle olivat muun muassa se, montako kertaa Loton potti on jäänyt tarkasteluhetkellä jakamatta sekä vuodenaika. Myyntien mallintamisessa onnistuttiin selkeästi parhaiten kertapelien osalta, yleisesti voidaan sanoa, että tulokset olivat sitä luotettavampia, mitä lyhyemmästä kestoista oli kyse.

Simulaattorin avulla Loton pelaamisesta havaittiin, että vuositasolla Loton myynti vaihtelee kymmeniä miljoonia euroja (suuruusluokka 10 % vuositason myynnistä) riippuen siitä, millaisia potteja vuoden ajalle sattuu osumaan. Tätä havaintoa voidaan pitää ehdottomasti yhtenä tämän työn mielenkiintoisimmista tuloksista. Sattuman vaikutuksen lisäksi havaittiin, että vuositason pelaamiseen vaikuttaa hyvin voimakkaasti myös se, millaista pottitarjoamaa Lotossa käytetään.

Myyntien mallintamisen ja simulointien avulla saatuja tuloksia voidaan pääsääntöisesti pitää luotettavina. Niitä tulkittaessa tulee kuitenkin huomata, että osaa selittäjistä ei malleihin kyetty sisällyttämään ja toisaalta osa selittäjäkandidaateista huomattiin vasta, kun tulokset oli jo laadittu. Vaikka Loton pelaamisen vaikuttavien selittäjien voidaan olettaa muuttuvan suhteellisen hitaasti ajassa, saattaisi olla mielenkiintoista tutkia, kuinka tulokset muuttuisivat, jos niissä huomioitain myös uudistuneen Loton pelikierrokset.

Alkusanat ja kiitokset

Työn valmistumiseen meni aikaa huomattavasti enemmän kuin alun perin oli tarkoitus. Näin ollen osa tuloksista mahdollisesti muuttuisi, jos vastaava tutkimus laadittaisiin nyt uudelleen. En usko että Loton pelaamiseen vaikuttavat tekijät ovat juuri muuttuneet, niiden vaikuttavuuksien suuruusluokassa on kuitenkin saattanut muutos tapahtua.

Työn valmistuminen on ollut monilta osin erittäin mielenkiintoinen prosessi, mutta toisaalta on hienoa saada se myös päätökseen. Työn aikana olen oppinut Lotosta ja sen pelaamisesta ylipäättään erittäin paljon. Itse pidän sitä työn suurimpana antina, sen lisäksi että kehityin merkittävästi matemaattisten mallinnosten ja simulointien laatimisessa.

Tahtoisin kiittää Veikkaus Oy:tä mielenkiintoisesta tehtävänannosta ja siitä, että olen saanut sieltä paljon olennaista tietotaitoa, jota ilman tämä työ ei olisi koskaan valmistunut.

Veikkaukselta haluaisin kiittää erityisesti ohjaajaani Harri Järvistä, joka varsinkin työtä aloitettaessa antoi monia vinkkejä siihen, mihin asioihin kannattaisi tarkemmin perehtyä ja mitkä olisi syytä jättää vähemmälle huomiolle. Haluaisin myös kiittää kaikkia muita Veikkaus Oy:n henkilökunnasta, jotka ovat työn valmistusta omalla panoksellaan edesauttaneet.

Kiitokset ovat paikallaan myös prof. Ehtamon suuntaan, joka jaksoi odottaa työn valmistumista.

Vantaa 22.1.2015

Ville Venojärvi

SISÄLLYSLUETTELO

1	Johdanto.....	1
1.1	Taustaa	1
1.2	Työn tavoitteista.....	1
1.3	Työssä käytetyistä menetelmistä	2
1.4	Työn rakenne	2
2	Loton lakiteknisestä perustasta sekä matemaattisesta rakenteesta	3
2.1	Loton vaatimukset laissa ja asetuksissa	3
2.2	Lotto Suomessa vuonna 2012 – rahapeli näkökulma	4
2.3	Loton rakenne ja voittojen määrittäminen	5
2.4	Loton vaikeusaste ja voiton odotusarvo.....	6
2.5	Loton myynti ja vaihto	8
2.6	Lotto ja Veikkaus Oy:n voittorahasto.....	9
3	Lottopelien historiasta	11
3.1	Lottopelien Eurooppalaista historiaa.....	11
3.2	Loton suomalaista historiaa	12
3.2.1	Loton syntymisen taustoista.....	12
3.2.2	Loton historian merkittäviä tapahtumia.....	14
3.2.3	Loton myynnin kehittyminen.....	16
3.2.4	Loton merkitys Veikkauksen liiketoiminnalle.....	17
4	Lottopelien tutkimuksista.....	18
4.1	Lottopelien tutkimusten tavoitteista, menetelmistä ja tuloksista	18
4.2	Ajatuksia lottopelien tutkimuksista	26
5	Loton myynnin mallintamisen menetelmä ja simulaattorin toimintaperiaate	27
5.1	Usean selittäjän regressiomallin rakenne, oletukset ja parametrien estimointi.....	28
5.2	Usean selittäjän regressiomallin selittäjien identifiointi, mallin diagnostinen testaus ja sillä ennustaminen.....	29
5.3	Simulaattorin tavoitteet ja toimintaperiaate	30
6	Mallinnoksissa käytetyn aineiston kuvaus.....	31
6.1	Tilastollisen aineisto ja sen tulkintaa.....	31
7	Loton myynnin mallintamisen tulokset.....	33
7.1	Loton myynnin mallien identifiointi	34
7.2	Loton myynnin mallintamisen tulokset.....	34
7.3	Regressiomallien grafiikkaa ja diagnostinen testaus	43
7.3.1	Regressiomallien kuvaajia	44

7.3.2	Diagnostinen testaus.....	46
7.3.2.1	Poikkeavat havainnot ja regressiokertoimien vakioisuus	46
7.3.2.2	Selittäjien multikollineaarisuus	49
7.3.2.3	Residuaalien homoskedastisuus, normaalius ja autokorrelaatio.....	50
7.3.2.4	Regressiomallien ennustuskyky ja sen tilastollinen testaus.....	54
7.4	Loton myynnin mallintamisen haasteita ja jatkokehitys mahdollisuuksia.....	58
8	Loton pelimyynnin simulointi.....	59
8.1	Simulointimallin kuvaus.....	59
8.2	Simulointitulokset.....	62
8.3	Loton pottien vaikutus myyntiin ja sen muuttuminen ajassa	72
8.4	Simulointimallin uskottavuus tarkasteluja sekä kehitysmahdollisuuksia	76
9	Johtopäätökset.....	76
	Lähdeluettelo	79
	Liitteet.....	81

Lyhenteet ja termistö

Parimutuaali: Tämän työn yhteydessä parimutuaali tarkoittaa rahapelimuotoa, jossa pelaajille palautettava osuus on vakio pelikierroksesta toiseen. Pelaajien voitot määräytyvät pelin sääntöjen mukaan. Voittoluokille allokoitava jakosumma on yleensä kiinteä prosenttiosuus pelin vaihdosta.

Rahapeli: Peli, jossa vastiketta (rahaa) vastaan osallistutaan rahapalkintojen arvontaan.

Kierrosmyynti (myynti): Loton kokonaismyynti yhden pelikierroksen aikana. Tarkasteluajankohta alkaa yksittäisen arvonnän suoritushetkestä ja päättyy seuraavan arvonnän suoritukseen. Tavanomaisesti Lotossa kierrosmyynti koostuu sunnuntain (edellinen arvonta lauantaina) ja lauantain välisestä kokonaismyynnistä.

Kokonaismyynti: Loton kaikkien eri kestojen (1,2,3,5,10 viikkoa) yhteenlaskettu euromääräinen myyntiarvo.

Kierrosvaihto (vaihto): Kuvaa Loton yksittäisen pelikierroksen arvontaan osallistuvien pelirivien euromääräistä kokonaisarvoa.

Pelikierros (kierros): Loton osalta voidaan määritellä, että pelikierros päättyy, kun arvontaan ei voi enää osallistua. Vastaavasti samalla alkaa seuraava pelikierros.

Voittorahasto: Tässä työssä voittorahasto-termiä käytetään kuvaamaan sekä Veikkaus Oy:n maksamattomien pelitoiminnan varoja että Loton lakisääteistä 2 % osuutta. Voittorahasto sisältää edellä kuvatut varat yhteenlaskettuna. Maksamattomat pelitoiminnan varat koostuvat Veikkaus Oy:n rahapeliin lunastamattomista voitoista, voittojen pyörityksistä sekä voitto-osuuksien tasauksista. Maksamattomat pelitoiminnan varat ja niiden kertyminen on määritelty valtioneuvoston asetuksessa 1288/2011 koskien Veikkaus Oy:n rahapeliin toimeenpanoa. Loton kunkin kierroksen kokonaisvaihdosta siirretään 2 % (Loton kierrokseen 40/2011 saakka 0,9 %) voittorahastoon, ja nämä varat maksetaan Loton voittoina pelin myöhemmillä kierroksilla.

Simulaattori: Tässä simulaattorilla tarkoitetaan matemaattista ohjelmistoa, jonka avulla Loton arvontoja suoritetaan ja vuosittaisen myyntitasoja estimoidaan.

Pottijakauma: Jakauma eri pottityypeistä tarkasteltavana ajankohtana.

Pelikate: Prosenttiosuus, jonka pelinjärjestäjä pitää itsellään pelin vaihdosta. Jos esimerkiksi pelin palautusprosentti on 40%, niin pelikate on tällöin 60%.

Rollover: Tässä työssä rollover kuvaa tilannetta, jossa Lotossa ei löydy päävoittoja ja päävoittoluokan osuus siirretään maksettavaksi päävoitoille pelin myöhemmillä kierroksilla.

Kokonaisjakosumma: Kierroskohtainen kokonaisjakosumma (jakosumma) tarkoittaa pelin voittoluokkien sääntöjen mukaan määrättyjä euromääräisiä osuuksia pelin voitonjaosta. Kokonaisjakosumma koostuu eri voittoluokkien yhteenlasketuista jakosummista.

Pottityyppi: Kuvaa sitä, kuinka moneen kierrokseen ei ole löytynyt päävoittoa. Jos esimerkiksi päävoittoa ei ole löytynyt kahteen edelliseen kierrokseen, niin kuluvaan kierrokseen pottityyppi on kaksi.

Ylimääräinen potti: Kaikkien varsinaiseen lottoarvontaan osallistuneiden pelirivien kesken arvottu ylimääräinen potti, joka jaetaan seitsemän oikein -tulosten kesken.

Peruspotti: Pottityypin nolla omaavat potit (kierrokset).

Kertapeli: Yhden viikon kestona pelatut pelirivit.

Kestopeli: Kahden, kolmen, viiden tai kymmenen viikon kestona pelatut pelirivit.

Pottitarjoama: Loton ehdollinen pottiputki, mikäli päävoittoa ei löydy. Pottitarjoama määrittää sen, kuinka suuri potti milläkin pottityypillä pelissä on tarjolla. Esimerkiksi: 1,1-3-5-7-9-11M€ kuvaa sitä että peruspotit ovat 1,1M€, seuraavaksi mikäli päävoittoa ei löydy tarjolla on 3M€ ja niin edelleen.

Päävoitto (Potti): Loton seitsemän oikein tuloksille tarjolla oleva voittosumma.

Voittoluokka: Kuvaa ehdot yksittäiselle tulokselle, jolla pelaaja on oikeutettu voittoon pelissä. Esimerkiksi Lotossa ensimmäinen voittoluokka on seitsemän oikein ja toinen voittoluokka on kuusi ja yksi lisänumero oikein.

1 JOHDANTO

1.1 TAUSTAA

Lottoa on pelattu Suomessa vuodesta 1971 lähtien ja tammikuussa 2011 Lotto täytti jo kunnioitettavat 40 vuotta. Lotto on ollut Veikkauksen suosituin rahapeli aina lanseerausvuodestaan saakka ja se omaa erittäin vahvat perinteet suomalaisessa rahapelikulttuurissa.

Lottopelit ovat maailmanlaajuisestikin suosittuja rahapelejä. Loton tyyppisiä pelejä tarjotaan kansallisten rahapeliyhtiöiden toimesta niin Euroopassa, kuin sen ulkopuolellakin.

Lottopeleille tyypillisiä piirteitä ovat muun muassa mahdollisuus suureen päävoittoon, suhteellisen alhainen osallistumismaksu ja toisaalta myös päävoiton epätodennäköisyys. Lottopelien viehätys perustuukin suurelta osin siihen, että kohtalaisen pienellä panostuksella pelaajat saavat mahdollisuuden haaveilla siitä, mitä kaikkea tekisivätkään päävoiton osuessa kohdalle.

1.2 TYÖN TAVOITTEISTA

Yleisesti ottaen tämän työn ensisijaisena tavoitteena on ollut perehdyttää kirjoittajaa Loton rahapeliominaisuuksiin sekä niihin seikkoihin, jotka Loton pelaamiseen vaikuttavat. Työn matemaattisen osion, eli mallintamisen ja simuloinnin avulla, tekijä on päässyt tutustumaan niin matemaattisten mallinnusmenetelmien kuin Monte Carlo -simuloinnin hyödyntämiseen liiketoiminnallisessa ympäristössä.

Työn eräänä osatavoitteena on ollut mallintaa Loton kierroskohtaista myyntiä ja tutkia, mitkä tekijät siihen vaikuttavat. Erityisenä mielenkiinnonkohteena on ollut se, kuinka erilaiset pottitarjoamat vaikuttavat pelaamiskäyttäytymiseen. Työn keskeisin tavoite on ollut mallinnosten avulla muodostaa matemaattinen ohjelmisto, jonka avulla voidaan simuloida muun muassa Loton vuositason myyntijakaumia. Simulaattorin avulla on eritoten tarkoitus tutkia sitä, kuinka erilaiset pottitarjoamat pelaamiseen vaikuttavat ja millaisia voittorahaston kulutuksia voidaan milläkin pottitarjoamamalleilla odottaa.

Mallintamis- ja simulointiosion lisäksi työssä on tarkasteltu lyhyesti Loton historiaa niin kotimaassa, kuin muualla Euroopassakin. Työssä on paneuduttu myös Loton ominaisuuksiin rahapelinä ja tarkasteltu esimerkiksi nykyisen Loton voittojen suuruuksien määrittämisen vaiheita. Edellisten lisäksi on myös tehty kevyt katsaus Loton lakitekniisiin ominaisuuksiin.

Työtä tehdessä on käynyt selväksi, kuinka haasteellista Loton myynnin mallintaminen on. Vaikka muodostettujen regressiomallien voidaan katsoa olevan kohtuullisen hyviä, on olemassa selkeästi vielä myyntiin vaikuttavia tekijöitä, joita tässä tehdyt mallit eivät sisällä. Tätäkin oivallusta voidaan pitää tämän oppimisprosessin eräistä keskeisimmistä anneista.

1.3 TYÖSSÄ KÄYTETYISTÄ MENETELMISTÄ

Loton myynnin mallintaminen on toteutettu kahden toisistaan poikkeavan usean selittäjän regressiomallin avulla. Toinen malleista on tavanomainen usean selittäjän regressiomalli, jonka lisäksi laadittiin myös logaritmoitu regressiomalli. Vaihtoehtoisina menetelminä olisi ollut esimerkiksi erilaisten aikasarjojen tai neuroverkkojen hyödyntäminen. Työtä tehdessä päädyin kuitenkin usean selittäjän regressiomallien hyödyntämiseen, sillä ne suoriutuivat järjestään mallinnusprosessista yhtä hyvin tai paremmin kuin esimerkiksi SARIMAX- mallit. Neuroverkkojen hyödyntämisen hylkäsin lopulta sen takia, että tahtotilana oli myös tulkita sitä, miten mikäkin selittäjä myyntiin vaikuttaa. Neuroverkkojen kaltainen "black box" -malli ei olisi tässä suhteessa ollut relevantti. Usean selittäjän regressiomallien valintaa tuki myös se, että lottopelien mallintamista käsittelevissä tieteellisissä julkaisuissa on usein menetelmänä hyödynnetty juuri usean selittäjän regressiomallia.

Työssäni päädyin erottelemaan Loton myynnin kestojen mukaan. Näin ollen jokaiselle kestolle muodostettiin oma regressiomalli. Malleiksi muodostui siis: yhden, kahden, kolmen, viiden ja kymmenen viikon myyntien mallit. Tämä mahdollisti esimerkiksi Loton potin vaikutuksen tutkimisen eripituisten kestojen pelaamiseen. Eräänä mahdollisuutena olisi myös ollut kuvata Loton vaihtoa yhtenä kokonaisuutena, joka sisältää kerta- ja kestopelit. Tässä toteutustavassa regressiomalleja olisi tullut ainoastaan yksi, eikä jokaiselle kestolle omaansa. Sen heikkoutena voidaan kuitenkin pitää esimerkiksi sitä, ettei malli olisi kyennyt yhtä tehokkaasti selittämään kestopelaamisen muuttumista erilaisilla pottitasoilla.

1.4 TYÖN RAKENNE

Työn rakenne voidaan jakaa karkeasti neljään eri osioon. Ensimmäisessä osiossa käsitellään johdannon lisäksi Loton lakiteknistä perustaa sekä Loton matemaattista rakennetta rahapelinä (kappale 2). Ensimmäiseen osioon voidaan katsoa kuuluvan myös lottopelien historiaa käsittelevä kappale 3, jossa on kuvattu niin suomalaisen Loton kuin lottopelien historiaa yleisesti.

Työn toisen osion voidaan katsoa sisältävän kappaleet 4-6. Kappaleessa 4 on käsitelty lottopelien pelaamiseen liittyviä tieteellisiä tutkimuksia sekä niiden tuloksia. Kappaleessa 5 on kuvattu tarkemmin niitä matemaattisia menetelmiä, joita on tässä työssä Loton pelaamisen tutkimiseen hyödynnetty. Kappaleessa 6 on puolestaan

esitetty tilastollisten analyysien tuloksia Loton myynnin mallintamisessa käytetystä aineistosta.

Työn kolmannessa osiossa (kappaleet 7 ja 8) on esitetty Loton myynnin mallintamisen sekä simulointien tulokset. Kappaleessa 7 on kuvattu Loton myynnin mallintamisen tulokset niin tavanomaisen kuin logistisen regression menetelmillä. Kappaleessa 8 puolestaan on kuvattu kappaleessa 7 muodostettujen regressiomallien avulla laaditut varsinaiset simulointitulokset.

Työn lopuksi on esitetty tärkeimmät johtopäätökset ja opit kappaleessa 9.

2 LOTON LAKITEKNISESTÄ PERUSTASTA SEKÄ MATEMAATTISESTA RAKENTEESTA

2.1 LOTON VAATIMUKSET LAISSA JA ASETUKSISSA

Vuoden 2012 alussa Suomessa astui voimaan rahapelejä koskeva legaalinen monopoli. Legaalilla monopolilla rahapelien osalta Suomessa tarkoitetaan sitä, että rahapelejä saavat toimeenpanna Veikkaus Oy, Raha-automaattiyhdistys sekä Fintoto Oy. Alle olen lyhyesti poiminut Loton kannalta olennaisia kohtia niin voimassa olevasta arpa-jaislaista kuin valtioneuvoston asetuksesta Veikkauksen rahapeleistä. Osion lopussa on myös viittaus Loton sääntöihin.

VALTIONEUVOSTON ASETUS VEIKKAUS OY:N RAHAPELIEN TOIMEENPANOSTA (1288/2011)

1 §

Peleihin osallistuminen

"Veikkaus Oy:n, jäljempänä yhtiö, toimeenpanemiin rahapeleihin osallistutaan maksamalla pelimerkintöjä vastaava maksu tai ostamalla arpa. Pelimerkinnät toimitetaan yhtiöön tietoliikenneyhteyksiä käyttäen, teknisenä tallenteena tai kuponkitietoina. Pelaaminen voi tapahtua joko asiamiespisteissä tai itsepalveluna puhelimen, tietokoneen taikka niihin verrattavan teknisen sovelluksen välityksellä."

3 §

Pelimaksuista pelaajille voittoina maksettavat osuudet

"5) muissa veikkauksissa keskimäärin 40–55 prosenttia kokonaisvaihdesta;"

ARPAJAISLAKI 23.11.2001/1047 & LAKI ARPAJAISLAIN MUUTTAMISESTA 575/2011

2 §

Arpajaisten määritelmä

"Tässä laissa tarkoitetaan arpajaisilla toimintaa, johon osallistutaan vastiketta vastaan ja jossa osallistuja voi saada kokonaan tai osittain sattumaan perustuvan rahatarvoisen voiton. ..."

3 §

Arpajaisten toimeenpanomuotojen määritelmät

"3) veikkauspeleillä muita kuin 2 kohdassa tarkoitettuja arpajaisia, joissa pelaajalla on mahdollisuus osallistua urheilukilpailujen tuloksiin taikka numeroiden tai merkkien arpomiseen liittyvien arvausten perusteella pelimaksuista kertyvien voittojen jakoon;"

Loton säännöt

Kohdat 1.-36. Ks. Liitteet.

Edellä esitetyt kohdat niin rahapeliasetuksesta, arpajaislaista kuin Loton säännöistäkin ovat olennainen osa sitä lakiteknistä kokonaisuutta, joka mahdollistaa Loton toimeenpanemisen Suomessa Veikkauksen toimesta.

2.2 LOTTO SUOMESSA VUONNA 2012 – RAHAPELI NÄKÖKULMA

Harri Järvinen määritteli diplomityössään (2007) rahapelin seuraavasti (lainauksesta on poistettu otteita soveltuvien osien):

"Perinteiset rahapelit eli mm. kaikki Veikkauksen pelit arpoja lukuun ottamatta rakentuvat neljästä sääntöjen määrittelemästä pääpalikasta: arvonnasta, pelaajan käyttöliittymästä, voitonjakoista ja ... vertailusta. Lyhyesti: arvonnasta tulosta ja pelaajan jättämää riviä verrataan toisiinsa ja vertailun tulosta edelleen verrataan ennalta määritettyyn voitonjakoon, josta selviää pelaajan rivin kerroin. Pelaajan voitto on tämän voitonjaon kertoimen ja pelaajan panoksen tulo."

Suomessa Veikkauksen toimeenpanema Lotto täyttää edellä esitetyt neljä kriteeriä rahapelille

- Lotossa suoritetaan arvonta. Arvonta tapahtuu tyypillisesti lauantai-iltaisina (poikkeuksellisesti sunnuntai-iltaisina) ja siinä arvotaan 7 numeroa ja kaksi lisänumeroa kolmestakymmenestä yhdeksästä. Arvonnassa rakenne on aina sama (ellei pelin sääntöjä muuteta)
- Pelaajan mahdollisia käyttöliittymiä ovat esimerkiksi pelikuponki (ks. Liitteet kuva 9) ja Veikkauksen verkkopalvelu veikkaus.fi.

- Loton voitonjako on määritelty sääntökohdissa 20.-29. (Ks. Liitteet, Loton säännöt)
- Vertailu pelaajan pelirivien ja arvonnän tuloksen välillä suoritetaan arvontatuloksen ratkettua. Vertailun tuloksena selviää mahdollinen voitto ja sen suuruus

Loton voidaan katsoa täyttävän edellä esitetyt rahapelin määritelmät (sekä lakitekniset vaatimukset kohdasta 2.1). Näin ollen Lottoa voidaan tästä eteenpäin tarkastella Suomessa laillisesti toimeenpantavana rahapelinä.

2.3 LOTON RAKENNE JA VOITTOJEN MÄÄRITTÄMINEN

Seuraavassa määrittelen lyhyesti 41/2011 alkaen toimeenpantun Loton rahapelirakenteen. Osiossa on myös esitetty esimerkinomaisesti Loton voittojen suuruuksien määrittäminen.

- Lotossa suoritetaan yksi arvonta, jossa arvotaan seitsemän (7) numeroa ja kaksi (2) lisänumeroa kolmestakymmenestäyhdeksästä (39). Arvonta suoritetaan palauttamatta (arvottuja palloja ei palauteta rumpuun kesken arvonnän)
- Lotto on parimutuaali rahapeli, eli kiinteä % -osuus pelin vaihdosta siirretään palautettavaksi pelaajille. Pelaajille palautettava osuus on sama kierroksesta toiseen
- Pelin kierrosvaihdosta (vaihto) pelaajille palautetaan 39,1%. Tämä muodostaa pelin kierroskohtaisen kokonaisjakosumman (jakosumma). Tämän lisäksi pelin vaihdosta siirretään 2,0% jaettavaksi pelin myöhemmillä kierroksilla. Näin ollen pelin vaihdosta pelaajille palautetaan $39,1\% + 2,0\% = 41,1\%$.
- Kierrosvaihdon jakautuminen voittoluokkien kesken:
 - Päävoittoluokan osuus vaihdosta on 16,5%
 - Kierroskohtaisesta jakosummasta maksetaan pelin kiinteät voitot (voittoluokat 10. ja 11.)
 - Jäljelle jäänyt jakosumma (kun päävoiton osuus ja kiinteät voitot on maksettu) jaetaan voittoluokkien 2.-9. kesken pelin säännöissä kuvatulla tavalla

Taulukossa 1 on kuvattu esimerkinomaisesti, kuinka Loton voitonjako lasketaan ja kuinka kierroskohtaisen jakosumman jakautuminen voittoluokille määritetään. Esimerkkierroksena on 21.4.2012 arvottu Loton kierros 16/2012.

Voittojen suuruuksien laskentaprosessi lyhyesti:

- 1) Kierrosvaihdosta lasketaan kierroksen kokonaisjakosumma (39,10 % vaihdosta). Esimerkissä: 3 375 598,40€
- 2) Kokonaisjakosummasta vähennetään 7 oikein tulosten jakosumma (16,50% vaihdosta). Esimerkissä: 1 424 485,26€

- 3) Kokonaisjakosummasta vähennetään kiinteiden voittojen jakosumat (tulokset 3+2 ja 3+1). Esimerkissä 50 200,00€ ja 169 805,00€
- 4) Jäljelle jäänyt jakosumma jaetaan voittoluokkien 2.-9. kesken taulukossa esitettyjen prosenttiosuuksien mukaan
- 5) Kun kierroskohtainen jakosumma on voittoluokille jaettu, määritetään voittojen lukumäärät voittoluokittain. Mikäli voittoja ei löydy jossain voittoluokassa toimitaan Loton sääntöjen edellyttämällä tavalla (Ks. Liitteet: Loton säännöt 20.-29.)
- 6) Suoritetaan mahdollinen voittoluokkien jakosummien täydennys, tämä edellyttää että voittoluokassa on löytynyt voittajia. Esimerkissä 7 oikein tuloksen jakosummaa täydennetään 75 514,74€:lla. Esimerkissä täydennys suoritetaan Veikkaus Oy:n voittorahastosta. Lotossa on viime vuosina periaatteena pidetty, että mainostettua jättipottia pienempää päävoittoa ei pelaajille makseta, vaan tarvittaessa jakosummaa täydennetään vastaamaan tarjolla ollutta jättipottia
- 7) Määritetään voittoluokkien bruttovoittojen suuruudet (jakosumma lisättynä mahdollinen jakosumman täydennys)/voittajien lkm.
- 8) Bruttovoitot pyöristetään alempaan kymmeneen senttiin. Erotus siirretään voittorahastoon rahapeliasetuksen vaatimusten mukaisesti (Valtioneuvoston asetus Veikkaus Oy:n rahapeliin toimeenpanosta 1288/2011). Taulukkoa tarkasteltaessa tulee huomioida että voittorahastoon siirrettävää summaa laskettaessa käytetään pyöristämättömiä lukuarvoja
- 9) Nettovoittojen suuruudet on saatu määritettyä ja voitot ovat valmiita maksumun

		Kierrosvaihto €		8 633 244,00				
VL	Tulos	Jako %	Jakosumma	Voittoihin lis.	Voittojen lkm.	Voitto brutto	Voittorahastoon	Voitto netto
1.	7	16,50 %	1 424 485,26	75 514,74	1	1 500 000,00	0,00	1 500 000,00
2.	6+1	8,80 %	152 337,52	0,00	10	15 233,75	0,05	15 233,70
3.	6	10,00 %	173 110,81	0,00	112	1 545,63	0,03	1 545,60
4.	5+2	2,00 %	34 622,16	0,00	12	2 885,18	0,08	2 885,10
5.	5+1	4,60 %	79 630,97	0,00	833	95,60	0,09	95,50
6.	5	13,50 %	233 699,60	0,00	4 611	50,68	0,08	50,60
7.	4+2	2,70 %	46 739,92	0,00	672	69,55	0,05	69,50
8.	4+1	16,80 %	290 826,17	0,00	17 948	16,20	0,00	16,20
9.	4	41,60 %	720 140,99	0,00	79 347	9,08	0,07	9,00
10.	3+2	5,00 €	50 200,00	0,00	10 040	5,00	0,00	5,00
11.	3+1	1,00 €	169 805,00	0,00	169 805	1,00	0,00	1,00
Yhteensä		39,10 %	3 375 598,40	75 514,74	283 391	3 451 113,14	6 590,04	3 444 523,10

Taulukko 1: Loton voitonjaon laskentaesimerkki

2.4 LOTON VAIKEUSASTE JA VOITON ODOTUSARVO

Loton vaikeusaste voidaan määrittää muun muassa kombinatoriikan tai hypergeometrisen jakauman avulla. Seuraavassa olen lyhyesti kuvannut yksittäisen Loton voittoluokan vaikeusasteen määrittämisen hypergeometrisen jakauman avulla

Muuttujat ja niiden arvot Lotossa

- $N = 39$ valittavissa olevien numeroiden lukumäärä
- $A = 7$ arvottavien numeroiden lukumäärä
- $L = 2$ arvottavien lisänumeroiden lukumäärä
- $R = 7$ pelaajan valitsema numeromäärä riviä kohden
- X satunnaismuuttuja joka kuvaa pelaajan oikeiden numeroiden lukumäärää
- Y satunnaismuuttuja joka kuvaa pelaajan oikeiden lisänumeroiden lukumäärää

Lotossa tuloksen $x + y$ oikein määrittäminen suoritetaan todennäköisyyslaskennan sääntöjen mukaan hypergeometrisen jakauman avulla seuraavasti:

$$P(x + y) = P(X = x) * P(Y = y|X = x) \quad (1)$$

Kaavan (1) ja hypergeometrisen jakauman nojalla saadaan:

$$P(x + y) = \frac{\binom{A}{x} \binom{N-A}{R-x}}{\binom{N}{R}} * \frac{\binom{L}{y} \binom{N-A-L}{R-x-y}}{\binom{N-A}{R-x}} \quad (2)$$

Edellä esitetty sievenee edelleen:

$$P(x + y) = \frac{\binom{A}{x} \binom{L}{y} \binom{N-A-L}{R-x-y}}{\binom{N}{R}} \quad (3)$$

Loton kunkin voittoluokan euromääräinen odotusarvo voidaan suoraviivaisesti laskea, kun tiedetään voittoluokalle allokoitava laskennallinen prosenttiosuus vaihdosta, voittoluokan vaikeusaste sekä pelin rivimaksu. Voittoluokan voiton suuruuden odotusarvo voidaan laskea kaavalla: (Järvinen; 2007)

$$E(\text{voitto}(x + y)) = h * \frac{S_{x+y}}{P_{x+y}} \quad (4)$$

Missä h kuvaa rivimaksua, S_{x+y} voittoluokalle $x + y$ vaihdosta laskennallisesti allokoitavaa prosentuaalista osuutta ja P_{x+y} voittoluokan $x + y$ todennäköisyyttä.

Taulukossa 2 on esitetty 41/2011 alkaen toimeenpannun Loton voittoluokkien laskennalliset jako-osuudet vaihdosta, todennäköisyydet, voittosuhteet (1/voittoluokan todennäköisyys), kombinaatiot eli montako eri riviä yksittäisessä arvonnassa sisältää kyseisen tuloksen, sekä voittoluokkien voittojen suuruuksien laskennalliset odotusarvot.

VL	Tulos	Jako %	Lask Jako-%	kombinaatioita	Todennäköisyys	Voittosuhde	E(voitto) €
1.	7	16,50 %	16,50 %	1	0,0000001	15 380 937	2 537 855
2.	6+1	8,80 %	1,78 %	14	0,0000009	1 098 638	19 585
3.	6	10,00 %	2,03 %	210	0,0000137	73 243	1 484
4.	5+2	2,00 %	0,41 %	21	0,0000014	732 426	2 967
5.	5+1	4,60 %	0,93 %	1 260	0,0000819	12 207	114
6.	5	13,50 %	2,73 %	9 135	0,0005939	1 684	46
7.	4+2	2,70 %	0,55 %	1 050	0,0000683	14 649	80
8.	4+1	16,80 %	3,40 %	30 450	0,0019797	505	17
9.	4	41,60 %	8,43 %	142 100	0,0092387	108	9
10.	3+2	5,00 €	0,49 %	15 225	0,0009899	1 010	5
11.	3+1	1,00 €	1,85 %	284 200	0,0184774	54	1
Yhteensä			39,10 %	483 666	0,03	31,8	

Taulukko 2: Loton voittoluokkien vaikeusasteet, sekä voittojen odotusarvot

2.5 LOTON MYYNTI JA VAIHTO

Tässä työssä käytetään usein termejä kierrosmyynti, myynti, kierrosvaihto ja vaihto. Näistä kierrosmyynti ja myynti tarkoittavat samaa asiaa, kuten myös kierrosvaihto ja vaihto. Vaihto ja myynti eroavat määritelmiltään toisistaan. Tämä ero johtuu siitä että Lotossa pelaajilla on mahdollisuus kestopelaamiseen (2,3,5,10 viikkoa) tavanomaisen yhden viikon, eli niin sanotun kertapelin lisäksi.

Myynti kuvaa sitä euromääräistä myyntiä, jolla tietyn pelikierroksen aikana Lottoa pelataan. Pelikierroksen aika voidaan määritellä seuraavasti: Yksittäinen pelikierros alkaa edellisen kierroksen arvonnän ajankohdasta ja päättyy tarkasteltavan kierroksen arvonnän suorittamiseen. Kierros kestää siis kahden perättäisen arvonnän välisen ajan. Tyypillisesti Loton kierrokset ovat sunnuntaista lauantaihin, sillä yleensä Loton arvonta suoritetaan lauantaisin.

Vaihto puolestaan tarkoittaa yksittäiseen arvontaan osallistuvien pelirivien euromääräistä arvoa. Vaihto voidaan määrittää kertomalla arvontaan osallistuvien pelirivien lukumäärä rivi hinnalla (poikkeustapauksena hinnankorotusta seuraavat kierrokset ja 10+1 kampanjat, jolloin arvunnoissa on mukana erihintaisia pelirivejä).

Alla on esitetty sekä myynnin että vaihdon määritelmät Lotossa. Kierroksen i euromääräinen myynti M_i eli kierroksen i aikana myytyjen Loton eri kestojen yhteenlaskettu arvo voidaan määrittää seuraavasti:

$$M_i = \sum_j m_{ij}; j = 1,2,3,5,10 \quad (5)$$

Kierroksen i euromääräinen vaihto V_i eli kierrokselle i osallistuvien rivien euromääräinen arvo määritetään vastaavasti:

$$V_i = \sum_j \sum_{k=i}^{i-j+1} \frac{m_{kj}}{j}; j = 1,2,3,5,10 \quad (6)$$

Edellä esitetyissä kaavoissa (5) & (6) m_{ij} kuvaa kierroksen i aikana myytyjen j mitaisten kestojen euromääräistä arvoa.

Loton yksittäistä kierrosta tarkasteltaessa voidaan vaihtoa pitää olennaisempana lukuna kuin myyntiä. Tämä perustuu siihen, että vaihto kuvaa arvontaan osallistuneiden pelirivien lukumäärää ja näin ollen se kertoo paremmin sen, kuinka paljon kyseinen Loton arvonta on pelaajia kiinnostanut. Tarkasteltaessa pidempää ajanjaksoa, esimerkiksi kalenterivuotta, ei ole merkitystä sillä, onko tarkastelun kohteena myynti vai vaihto, sillä ne ovat suuruusluokaltaan hyvin lähellä toisiaan.

Yksittäisen kierroksen osalta Loton myynti ja vaihto voivat erota toisistaan hyvinkin merkittävästi. Tällaisia tilanteita voi esiintyä esimerkiksi hinnankorotusten tai 10+1 kampanjoiden tapauksessa kun pelaajat pelaavat paljon pitkiä kestoja. Kuten osiossa 1.3 jo todettiin, tässä työssä päädyttiin mallintamaan eri kestojen myyntejä, joista sitten kaavan (6) avulla johdettiin kierrosvaihtojen estimaatit.

2.6 LOTTO JA VEIKKAUS OY:N VOITTORAHASTO

Veikkaus Oy:n maksamattomien pelitoiminnan varoihin kertyneet varat on käytettävä Veikkauksen eri rahapeliä voittoina valtioneuvoston asetuksen Veikkaus Oy:n rahapeliä toimeenpanosta (1288/2011) mukaan. Näitä varoja on maksettu historiassa myös Loton voittoina. Maksamattomien pelitoiminnan varojen lisäksi Loton kierroskohtaisesta vaihdosta siirretään 2 % pelin kierrosvaihdosta maksettavaksi Loton myöhemmillä pelikierroksilla. Pelikierrokseen 40/2011 saakka myöhemmille pelikierroksille siirrettävä osuus oli 0,9 % (tässä työssä mallinnoksissa käytettiin 0,9 %:n arvoa, joka vastasi mallinnostilannetta). Varoista, jotka muodostavat yhteenlaskettaessa maksamattomien pelitoiminnan varat ja Loton 2 % osuuden, käytetään tässä työssä nimeä voittorahasto.

Voittorahastoa on käytetty vuosien saatossa Loton voittoihin ainakin neljällä erilaisella tavalla. Voittorahaston avulla on maksettu suurempia päävoittoja, mitä tavanomainen jättipottikertymä olisi mahdollistanut. Tämän lisäksi Lotossa on järjestetty lisämaksuttomia ylimääräisen päävoiton arvontoja, joiden voitot on katettu voittorahastosta. Lotossa on myös aika ajoin ollut kampanjakierroksia, joiden aikana ostetut 10 viikon kestolotot ovat osallistuneet yhteentoista seuraavaan arvontaan (10+1kampanja). Edellä mainittujen lisäksi Loton voittoja on täydennetty hinnankorotusten yhteydessä vastaamaan tilannetta, jossa kukin pelirivi olisi tuonut Loton pelikierroksille vaihtoa korotetun riviannon verran. Seuraavassa on esitetty lyhyet esimerkit kustakin voittorahaston käyttötavasta.

Loton korotetut päävoitot

Tämä voittorahaston käyttötapa on viime vuosien aikana ollut kaikista yleisin. Viimeisten vuosien aikana Veikkauksen käytäntö Lotossa on ollut se, että Loton päävoi-

ton suuruus on aina ollut vähintään pelaajille kommunikoitavan potin verran. Tämä on johtanut siihen, että tarvittaessa päävoiton kattamiseksi on käytetty myös voittorahastoa. Esimerkkinä tällaisesta käyttötavasta voidaan pitää vaikka edellä jo kuvattua Loton kierrosta 16/2012, jolloin Loton pottina oli 1 500 000€. Kierroksen vaihto tuolla viikolla oli $8\,633\,244\text{€}$ ja näin ollen päävoiton osuudeksi muodostui $16,5\% * 8\,633\,244\text{€} = 1\,424\,485,26\text{€}$. Näin ollen päävoiton osalta katettavaa voittorahastosta jäi yhteensä 75 514,74€.

Aika ajoin Lotossa on myös pelaajille tarjottu niin sanottuja erikoispotteja, jolloin Loton päävoittona on ollut selkeästi suurempi summa kuin mitä pelkän päävoittokertymän nojalla olisi ollut. Esimerkkinä tästä voidaan pitää Loton kierrosta 38/2010, jolloin Lotossa juhlistettiin Veikkauksen 70-vuotisjuhlaa ja Loton pelaajille oli tarjolla tuohon mennessä Loton kaikkien aikojen suurin potti 7 200 000€. Kierrosvaihtona tuolloin oli $18\,743\,505,60\text{€}$ ja näin ollen päävoiton osuudeksi muodostui $13,6\% * 18\,743\,505,60 = 2\,549\,116,76\text{€}$. Kyseessä oli tuolloin pottityypin 0 pelikierros, joten katettavaa voittorahastosta jäi 4 650 883,24€.

Ylimääräisen päävoiton arvonnat eli lisäarvonnat

Lotossa on historian saatossa käytetty voittorahastoa järjestämällä ylimääräisen päävoiton arvontoja eli lisäarvontoja. Lisäarvonnat ovat olleet maksuttomia ja niihin on osallistuttu Loton varsinaisilla peliriveillä aina silloin, kun lisäarvonta on pelaajille ollut tarjolla. Lisäarvonnoissa on ollut ainoastaan yksi voittoluokka: seitsemän oikein.

Lisäarvonnoissa on ollut yleensä hieman pienempi potti kuin itse varsinaisessa Loton arvonnassa. Mikäli Lotossa on ollut pottina esimerkiksi 1 000 000€, on lisäarvonnassa puolestaan voinut olla tarjolla esimerkiksi 500 000€. Kierrosten 1/2006 - 10/2013 välisenä aikana Lotossa järjestettiin yhteensä 22 lisäarvontaa. Mikäli voittajia lisäarvonnoissa löytyy, niin pelaajien voitto katetaan kokonaisuudessaan voittorahaston avulla. Mikäli voittajia ei lisäarvonnassa löydy, niin tarjolla ollut summa siirtyy käytettäväksi varsinaisen Loton voittoina pelin myöhemmillä kierroksilla.

Loton lisäarvonnoilla on erilliset säännöt, joihin lukija voi halutessaan tutustua esimerkiksi Veikkauksen pelipalvelussa veikkaus.fi.

Ylimääräinen viikko 10 viikon kestolottoon eli 10+1 kampanja

Aika ajoin Lotossa on 10 viikon kestopeleille tarjottu kampanja, jossa pelaajien 10 viikon kestolotot ovat osallistuneet yhteentoista seuraavaan arvontaan. Pelaajat ovat maksaneet kestolotoistaan normaalin kymmenen viikon kestoloton hinnan. Tällöin pelaajien pelaamat kestot eivät ole tuoneet vaihtoa kampanjan yhdenteentoista arvontaan. Tämä tarkoittaa sitä, että yhdennentoista arvonnin euromääräinen vaihto on ollut alhaisempi kuin arvontaan osallistuneiden rivien mukaan laskettuna olisi ollut. Koska Loton voitot lasketaan euromääräisen vaihdon mukaan, johtaisi edellä ku-

vattu tilanne siihen, että yhdennellätoista kierroksella pelaajien voitot olisivat pienemmät kuin mitä ne olisivat tilanteessa, jossa 10+1 kampanjarivit eivät olisi arvontaan osallistuneet. Näin ollen yhdennellätoista kierroksella Loton voittoja täydennetään siten, että voitot vastaavat tilannetta, jossa jokainen arvontaan osallistunut pelirivi toisi kierrokselle vaihtoa normaalin rivimaksun verran.

Edellä kuvatut 10+1 kampanjat koskevat myös Loton yhteydessä pelattuja 10 viikon Lauantai-Jokereita, eli myös Lauantai-Jokereissa on vastaava voittojen täydennys yhdennellätoista kierroksella suoritettu. 10+1 kampanja oli pelaajille pelattavissa esimerkiksi kierroksella 48/2012 ja tällöin yhdenneksitoista kierrokseksi tuli 6/2013. Kierroksella 6/2013 voittorahastosta täydennettiin Loton voittoja 178 599,03€:lla ja Lauantai-Jokerin voittoja 34 223,28€:lla.

Loton voittojen täydennys hinnankorotusten yhteydessä

Mikäli Lotossa on korotettu hintaa, on pelaajilla ollut mahdollisuus osallistua vanhalla hinnalla Loton tuleviin arvontoihin. Vanhalla hinnalla tarkoitetaan tässä yhteydessä pelihetkellä voimassa olevaa hintaa. Tätä työtä laadittaessa edellinen Loton hinnan korotus astui voimaan kierroksella 41/2011. Mikäli pelaaja pelasi Lottoa esimerkiksi kymmenen viikon kestonä kierroksen 40/2011 aikana, maksoi hän kymmenen viikon kestopelistään sen summan mikä olisi pelihetken peliehtojen mukaisesti pelin hinnaksi muodostunut (0,80 €/pelirivi/kierros). Kierroksesta 41/2011 alkaen Loton rivi-hinta on ollut 1,00€/pelirivi/kierros. Kierroksien 41–49/2011 aikana Loton voitonjako suoritettiin siten, että voittojen suuruudet vastasivat tilannetta jossa kukin pelirivi, riippumatta siitä onko se pelattu uudella vai vanhalla hinnalla, olisi tuonut pelikierrokselle vaihtoa 1,00€/pelirivi. Voitonjaon täydennys suoritettiin voittorahastossa olevilla varoilla.

3 LOTTOPELIEN HISTORIASTA

3.1 LOTTOPELIEN EUROOPPALAISTA HISTORIAA

Lottopelien eurooppalaisen kehityksen lähteenä käytän teosta *Lotteries in Europe: Five centuries of history* (Bernard et. al.; 1994). Tässä työssä on kuvattuna kaksi esimerkkiä siitä, kuinka lottopelit ovat Euroopassa saaneet alkunsa. Myöhemmästä historiasta kiinnostuneet voivat tutustua esimerkiksi edellä mainittuun teokseen.

Lottopelejä aivan kuten muutakin raha-/uhkapelaamista on järjestetty jossain muodossa hyvin kauan. Eräitä ensimmäisiä lottopelejä oli Venetsiassa 1500 luvun alkupuolella järjestetyt arvonnat: *Venetsian Lotto*. Arvontojen perusluonne oli samanlainen kuin nykyisissäkin lottopeleissä: osallistumalla vastikkeelliseen arvontaan ja olemalla onnekas pelaajalla on ollut mahdollisuus voittaa hyvinkin arvokkaita palkintoja. Palkinnot olivat muun muassa hopeaa, arvokkaita koruja tai jopa arvokkaita

asuntoja. Jo tuohon aikaan lottopelien järjestämisen katsottiin olevan erinomainen keino kerätä varoja kansalta ja samalla vastata siihen luonnolliseen rahapelaamisen kysyntään, mikä oli jo tuolloin olemassa. Varoista hyötyivät hallinnollisten järjestelmien lisäksi muun muassa luostarit. Eräänä syynä Lottopelien menestykseen Venetsiassa voidaan pitää sitä, että kyseessä oli eräs merkittävimmistä Eurooppalaisista kauppapaikoista 1500-luvun alkupuolella. Näin ollen Venetsiassa liikkui rahan lisäksi paljon ihmisiä. Venetsian voidaan siis olettaa täyttäneen erään perusvaatimuksen, joka nykyisillään lottopeleillä on: potentiaalisen pelaajakunnan tulee olla riittävän laaja, jotta palkinnoista muodostuu pelaajien kannalta mielekkäitä. Edellä kuvattua voidaan pitää yhtenä menestyvän lottopelin vaatimuksista.

Toisena esimerkkinä varhaisesta ja menestyneestä lottopelistä on Genovassa järjestetyt arvonnat, joissa pelaajat veikkasivat sitä, ketkä viisi sadasta (sata vuodesta 1576 alkaen) ehdokkaasta arvottiin erilaisiin luottamustehtäviin (kolme paikkaa maaherrojen neuvostoon sekä kaksi prokuraattoria). Arvonnat kyseisille paikoille suoritettiin pelin alkuaikoina puolivuositain. Lottopelin arvontaa ei siis erikseen organisoitu, vaan hyödynnettiin tulosta, joka saatiin luottamushenkilöitä arvottaessa. Aluksi Genovan lottopeliä hallinnoi erinäiset yksityiset toimijat, mutta vuonna 1643 toiminta muuttui valtion monopolisoimaksi. Genovassa järjestetty lottopeli sai suurta suosiota myös muissa maissa ja monet maat, joissa vastaavien pelien järjestäminen oli kielletty, olivatkin huolissaan siitä, että merkittävästi varallisuutta siirtyisi pelin myötä Genovaan.

Yhteisenä piirteenä sekä Venetsian että Genovan peleille voidaan pitää sitä, että molemmat pelit saivat alkunsa aikana, jolloin molemmat sen aikaiset suuret kauppamahdit alkoivat menettää taloudellista asemaansa. Lottopelien avulla kauppakeskusten hallinnolliset elimet kykenivät hankkimaan varoja ja ylläpitämään yhteiskunnallista kehitystä tasolla, jolle ilman lottopelien tuottoja olisi ollut haastavaa päästä.

3.2 LOTON SUOMALAISTA HISTORIAA

Loton suomalaisen historian osuuden olen jakanut taustojen, merkittävien tapahtumien, pelaamisen historian ja merkityksen Veikkaukselle osioihin. Veikkauksen tietokantojen lisäksi lähteinä olen käyttänyt Heikki Ylikankaan vuonna 1993 julkaistua teosta: Markat Miljooniksi, sekä Kaija Kupiaisen diplomityötä: Lotto-pelin lanseeraus Suomeen (1971).

3.2.1 LOTON SYNTYMISEN TAUSTOISTA

Oy Tippaustoimisto Ab:n (nyk. Veikkaus Oy) ensimmäiset veikkaukset ratkottiin välirauhan päivinä 7.-8. syyskuuta vuonna 1940. Tästä sai alkunsa Suomessa laillisesti järjestettävä veikkaaminen, jonka tuotto on aina käytetty yleishyödyllisiin tarkoituksiin. Oy Tippaustoimisto Ab perustettiin 20:n silloisen johtavan urheilujärjestön toimesta ja veikkauksesta saatavat tuotot käytettiin aluksi kokonaisuudessaan urheilun

hyväksi. Opetusministeriö päätti, kuinka tuet lopulta jaettiin urheilujärjestöjen ja muiden organisaatioiden kesken. Vuoden 1950 aikana paine kuitenkin muiden edunsaajien mukaan ottamisesta kasvoi ja Suomen ooppera alkoikin saamaan tukea veikkauksvaroista, kiertoteitse tosin. Tässä en tarkemmin syvenny Veikkauksen voittovarojen jakamisen historialliseen kehitykseen, kiinnostuneen lukijan kannattaa tutustua edellä mainittuun Heikki Ylikankaan teokseen.

1960-luvun loppupuolella lotto alkoi nousemaan yhä varteenotettavammaksi pelimuotovaihtoehdoksi myös Suomessa, olihan Euroopassa jo lanseerattu menestyneitä lottopelejä. Veikkauksen sisällä ei kuitenkaan oltu yksimielisiä siitä, että lottoa pelimuotona kannattaisi Suomeen tuoda. Pelkona oli muun muassa se, että siihen asti suosituin pelimuoto vakioveikkaus kärsisi liian paljon uudesta, puhtaasti onneen perustuvasta pelistä. Näin ollen aluksi pyrittiinkin toteuttamaan Loton korvaavaa peliä urheilun avulla ja näin syntyikin tasapeliveikkaus. Pelistä ei kuitenkaan onnistuttu tekemään niin menestynyttä tuotetta, että se olisi poistanut paineet lottopelin lanseerauksesta.

Loton lanseerauksen eräänä huolena oli myös se, että ei ollut selkeää näkemystä siitä, mikä yhtiö peliä tulisi toimeenpanemaan. Veikkaus ei ollut mitenkään itsestään selvä vaihtoehto Loton järjestäjäksi. Syynä tähän on se, että tuolloiseen lainsäädäntöön ei ollut kirjattu Loton kaltaista peliä. Muun muassa silloinen raharpajaisyhdistys nähtiin yhtiön sisällä potentiaalisena pelin operoijana.

Lopulta kuitenkin ilmapiiri yhtiön sisällä kääntyi Lottoa kohtaan suopeammaksi. Tähän vaikutti muun muassa tieto siitä, että peli saataisiin Veikkauksen järjestämäksi. Toisaalta havaittiin että Sveitsissä, jota voitiin pitää Suomen kanssa samankaltaisena markkinana, lottopeli oli lanseerattu varsin menestyksekkäästi. Tuohon aikaan tiedostettiin myös se tosiasia, että Veikkaukselle asetettuja kasvutavoitteita oli äärettömän haastava toteuttaa, ellei pelikatalogiin saataisi jotain täysin uuden tyyppistä tuotetta. Näkemys yhtiön sisällä alkoi myös kallistua sille kannalle, että vaikka urheilupelaaminen Loton myötä laskisi, tulisi pelin lanseeraus kuitenkin olemaan yhtiölle kannattava toimenpide.

Kun Lotto oli päätetty Suomeen toteuttaa, piti vielä määrittää, millainen pelistä tulisi tehdä. Tärkeimpänä kysymyksenä oli pelin vaikeusasteen asettaminen oikean suuruiseksi. Kaija Kupiainen kirjoittaa diplomityössään että Suomessa tavoitteena oli joka viikkoinen suurvoittaja. Alussa peli haluttiin siis mitoittaa kutakuinkin niin, että arvontaan osallistuisi yhtä monta peliriviä mitä pelin vaikeusaste on.

Pelin pelaamisen suuruusluokkaa arvioitaessa tehtiin muun muassa asiakaskysely, jolla pyrittiin kartoittamaan, kuinka suuri osa silloisista Veikkauksen asiakkaista osallistuisi uuteen numeropeliin - Lottoon. Vastaukset olivat rohkaisevia, sillä yli 70% asiakkaista vastasi myöntävästi. Toisessa kyselyssä saatiin vielä rohkaisevampia tu-

loksia, kun lähes 90% vastaajista ilmoitti kokeilevansa peliä lähes tai melko varmasti. Peliformaatti valikoitui lopulta pelin vaikeusasteen ja Sveitsin esimerkin pohjalta. Ensimmäinen Lottokierros pelattiin formaatissa 40/(6+3), eli pelaajat rastivat kuusi numeroa 40 mahdollisesta ja pelissä arvottiin kuusi numeroa ja kolme lisänumeroa.

Vaikeusasteena päävoiton osalta pelillä tuolloin oli 1:3 838 380. Pelin hinta määrättiin asiakaskyselyn ja asiantuntija-arvioiden perusteella ja peli lanseerattiin 50 pennin ruudukkomaksulla. Tuolloin pidettiin tärkeänä sitä, että päävoittajia löytyisi keskimäärin yksi joka viikko. Jotta tämä toteutuisi, jokaiselle kierrokselle tulisi osallistua pelin vaikeusastetta vastaava määrä ruudukoita, eli noin 4 miljoonaa. Tämän seurauksesta pelin kierrosmyyntitavoitteeksi 50 pennin rivimaksulla asetettiin 2 miljoonaa markkaa kierrosta kohden.

Lotto sai Suomessa erittäin innostuneen vastaanoton ja uudesta pelistä tuli jo ensimmäisen vuotensa aikana Veikkauksen suosituin tuote, jota se on ollut aina tähän päivään saakka, vieläpä ylivoimaisesti.

3.2.2 LOTON HISTORIAN MERKITTÄVIÄ TAPAHTUMIA

Lotto on ollut Veikkauksen suosituin rahapeli alkuaajoistaan lähtien. Pelin ensimmäinen arvonta suoritettiin 3.1.1971 ja oikeaksi riviksi muodostui 1, 9, 18, 19, 36 ja 37. Lisänumeroiksi ensimmäisellä kierroksella arvottiin 5, 25 ja 27. Päävoittajia löytyi tuolloin yksi kappale ja voittosummaksi nykyrahassa muodostui noin 60 000€.

Taulukoissa 3 ja 4 on listattu Loton rivimaksun ja pelimuotojen muutokset kronologisessa järjestyksessä. Lähteenä on käytetty Veikkauksen sisäisiä tietopankkeja.

Kuten taulukoista nähdään on hinnanmuutosten osalta strategia ollut selkeä: Loton hintaa on aika-ajoin korotettu, kulloisenkin taloustilanteen mukaan.

Pelimuotoon on tehty kaksi merkittävää uudistusta lanseerauksen jälkeen. Merkittäviksi muutoksiksi tässä ei lasketa arvottavien lisänumeroiden määrän muuttamista, sillä tällöin päävoiton todennäköisyys ei muutu. Ensimmäinen suuri muutos peliformaatissa toteutettiin kierroksella 40/1980, jolloin pelaajat alkoivat rastimaan seitsemän numeroa kuuden sijaan. Myös valittavissa olevien numeroiden lukumäärää pudotettiin 40→37. Päävoiton todennäköisyys muuttui tuolloin: 1:3 838 380 → 1: 10 295 472. Jälkimmäinen merkittävä muutos formaatissa toteutettiin 2/1986, jolloin pelaajat valitsivat edelleen seitsemän numeroa, mutta valittavien numeroiden määrä kasvoi nykyiseen 37→39. Päävoiton vaikeusaste muuttui tuolloin 1: 10 295 472 → 1: 15 380 937. Loton vaikeusaste on edelleen sama, sillä pelaajat valitsevat edelleen seitsemän numeroa 39:stä.

Päävoiton vaikeusastetta muuttamalla voidaan vaikuttaa siihen, kuinka usein päävoitto jää jakamatta ja syntyy niin sanottu jättipottikierros. Jättipottikierrokset vaikuttavat pelaamiseen positiivisesti, mikä tullaan myöhemmin näkemään.

Kierros	Mk	€
1/1971	0,50	0,084
1/1975	1,00	0,168
26/1983	1,50	0,252
44/1988	2,00	0,366
10/1992	2,50	0,420
6/1995	3,00	0,504
8/1999	3,50	0,588
1/2002	-	0,60
7/2004	-	0,70
6/2008	-	0,80
41/2011	-	1,00

Taulukko 3: Loton rivimaksun (rivihinnan) muutokset

Kierros	Arvotut numerot	Voittoluokkien lkm.
1/1971	6+3/40	3
1/1972	6+2/40	4
40/1980	7+4/37	5
2/1986	7+0/39	4
17/1986	7+3/39	5
41/2011	7+2/39	11

Taulukko 4: Loton pelimuodon muutokset

Taulukoissa 3 ja 4 esitettyjen muutosten lisäksi on syytä mainita Loton lisävoittoluokat, jotka olivat pelattavissa kierroksilla 24/2009 – 40/2011. Lisävoittoluokat olivat maksullinen lisäominaisuus (0,20€/rivi), jonka avulla pelaaja osallistui kuuden erillisen voittoluokan arvontaan. Arvontatuloksena käytettiin varsinaista Loton arvontatulosta.

Loton ehkä suurin uudistus (jos ei lasketa edellä kuvattuja vaikeusasteeseen vaikuttaneita muutoksia) pelin lanseerauksen jälkeen koettiin kierroksella 41/2011. Tuolloin pelistä lakkautettiin erilliset lisävoittoluokat, pelin voittoluokkien lukumäärää kasvatettiin 5→11, päävoittoluokan osuutta kasvatettiin 13,6% → 16,5%:iin ja rivimaksu muutettiin 0,80€ → 1,00€:oon. Tietyssä mielessä pelin lisävoittoluokat sulautettiin itse pääpeliin eli Lottoon. Muutokset olivat suurehkoja pelaajille, jotka eivät olleet lisävoittoluokkia pelanneet. Pelaajat, jotka olivat lisävoittoluokat aikaisemmin pelanneet, eivät kovin merkittävää muutosta pelin logiikassa kokeneet.

Edellä kuvattujen muutosten lisäksi muun muassa päävoittoluokan osuutta on aika ajoin tarkistettu. Asiakkaille se on näkynyt muun muassa kasvaneina päävoittoina. Eräs tällainen päävoittoluokan osuuden korotus toteutettiin kierroksella 6/2008, jolloin päävoittoluokan osuus muutettiin 10 % → 13,6 %.

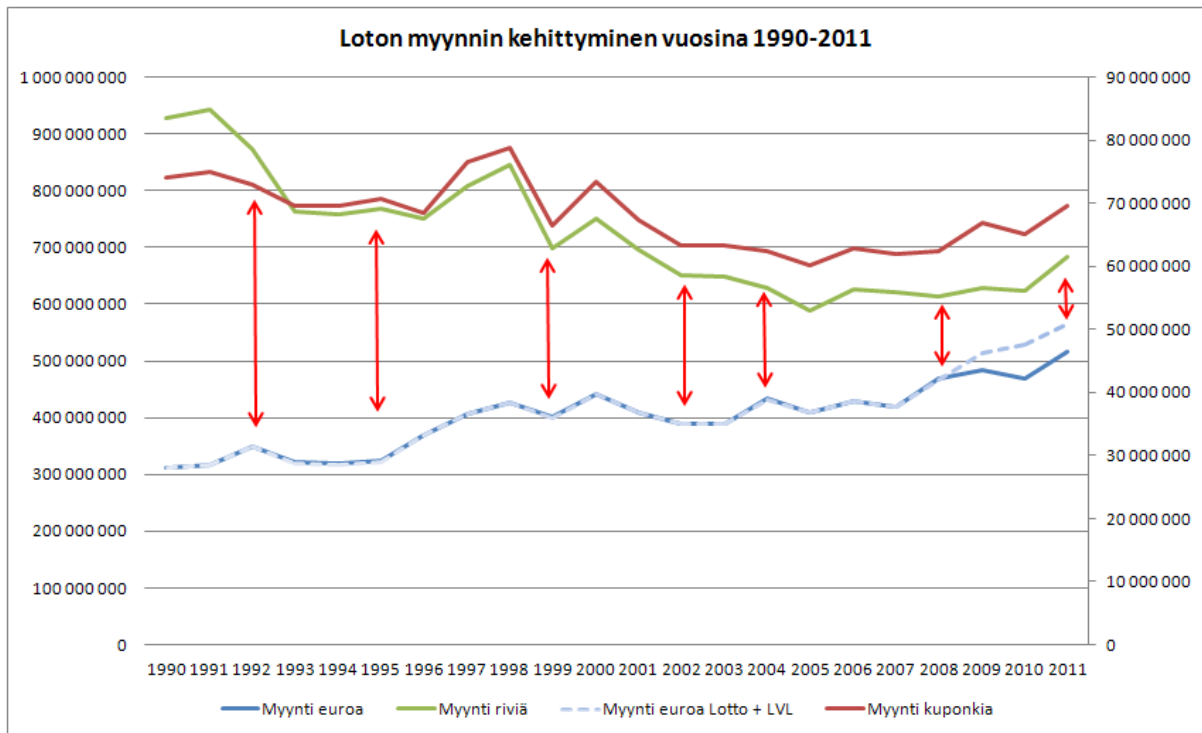
Edellä on kuvattu merkittävimpiä vuosien saatossa tehtyjä rakenteellisia muutoksia Lottoon. Muutoksista huolimatta pelin perusajatus on vuodesta toiseen kuitenkin säilynyt samana: pelaajat ovat pienellä panostuksella voineet haaveilla elämää muutavista voitoista.

3.2.3 LOTON MYYNNIN KEHITTYMINEN

Kuvassa 1 on esitetty Loton myynnin kehittyminen vuosien 1990-2011 välisenä aikana. Kuvasta on havaittavissa Loton suosion jatkuva kasvu euromääräisessä pelaamisessa. Punaiset nuolet osoittavat vuosia, joiden aikana Loton hintaa on korotettu.

Loton euromääräisessä pelaamisessa on vallinnut selkeä positiivinen trendi viimeisen parinkymmenen vuoden aikana. Joinain ajanjaksoina Loton pelaamisen kasvu on ollut maltillisempaa, mutta pääsääntöisesti suunta on ollut ylöspäin. Euromääräisesti tarkasteltuna voimakkaimman kasvun aikaa ovat olleet vuodet 1995-1998, sekä 2007-2011. Näinä ajanjaksoina jopa pelatut rivimäärät ovat kasvaneet hinnankorotuksista huolimatta. Osittain tällaisia hyviä jaksoja voidaan selittää ”hyvällä onnella”. Onnen merkitys Lotossa on siinä mielessä suuri, että kuten myöhemmin tullaan näkemään, niin ihmisten pelaamishalukkuuteen vaikuttaa voimakkaasti tarjolla oleva jättipotti. Mitä suurempia jättipotteja vuoden ajalle sattuu osumaan, sitä enemmän Lottoa odotusarvoisesti pelataan. Vuosien 2007-2011 positiivinen kehitys on entistä voimakkaampaa, jos Loton lisävoittoluokkien vaikutus huomioidaan. Kuvassa lisävoittoluokat on erotettu Loton myyntikäyrästä ylöspäin erkanevana janana.

Viimeisen vuosikymmenen aikana Loton kuponkimäärät ovat olleet hienoisessa kasvussa euromääräisen kasvun ohessa. Tästä voidaan suhteellisen luotettavasti päätellä että peli on säilyttänyt mielenkiintonsa laajan pelaajakunnan keskuudessa, eikä kasvanut pelaaminen ole keskittynyt ainoastaan harvojen pelaajien käsiin.

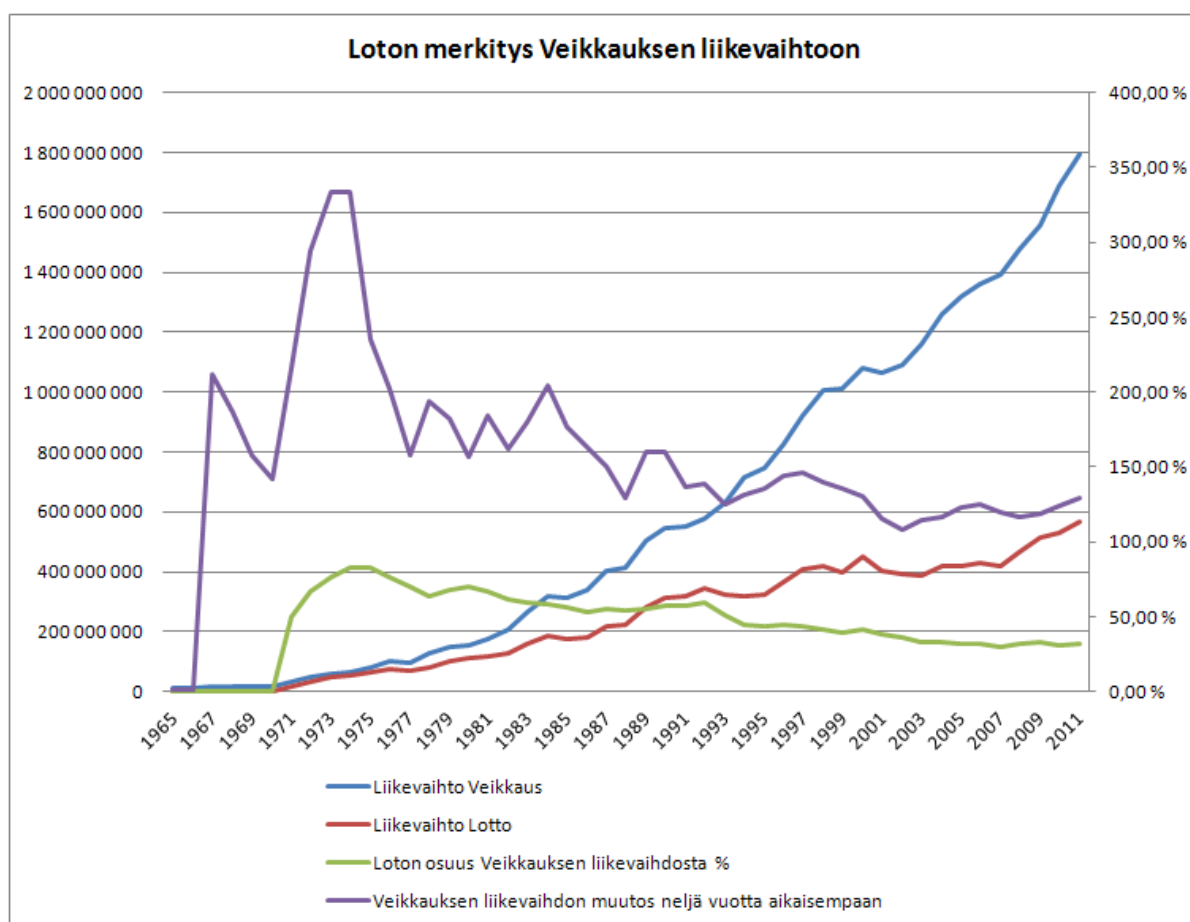


Kuva 1: Loton myynnin kehittyminen vuosina 1990-2011

3.2.4 LOTON MERKITYS VEIKKAUKSEN LIIKETOIMINNALLE

Kuvassa kaksi on esitetty Loton merkitys Veikkauksen liikevaihtoon vuodesta 1965 alkaen. Kuvassa esitettyjä euromääräisiä liikevaihtoja ei ole muutettu nykyrahaksi.

Kuvasta voidaan havaita Loton merkitys Veikkaukselle moneltakin kantilta. Parhaimmillaan Loton osuus Veikkauksen liikevaihdosta on ollut yli 80 prosenttia. Vaikka Loton merkitys liikevaihdossa yksittäisenä tuotteena on ajan mittaan laskenut, niin vielä vuonna 2011 Loton osuus oli yli 30 prosenttia Veikkauksen vuosivaihdosta. Jos tarkasteltaisiin vaihdon sijaan Veikkauksen pelikatetta, osuus olisi vielä huomattavasti suurempi, sillä Loton palautusprosentti on alhaisempi kuin valtaosan muista Veikkauksen rahapeleistä. Myös kuvaaja, joka näyttää Veikkauksen liikevaihdon muutosta neljän vuoden takaiseen, on hyvin mielenkiintoinen. Siitä voidaan nähdä että Loton lanseerauksen jälkeiset vuodet (1971-1975) Veikkaus kasvatti liikevaihtoaan suhteellisesti nopeimmin mitä kertaakaan vuoden 1965 jälkeen. Mikään muu tuote tai jakelukanavaudistus ei ole kyennyt vastaamaan Loton lanseerauksen jälkeen. Nykyaikana tuollaisten kasvulukujen saavuttaminen olisi tosin käytännössä mahdotonta, ottaen huomioon tämänhetkinen rahapelaamisen taso Veikkauksen peleissä ja ylipäättään Suomessa.



Kuva 2: Loton merkitys Veikkauksen liikevaihtoon

4 LOTTOPELIEN TUTKIMUKSISTA

Ottaen huomioon lottopelien kansainvälisesti merkittävä asema rahapelimarkkinoilla, on lottopelien myynnin mallinnoksia ja niiden tuloksia käsitelty myös tieteellisissä julkaisuissa. Löytämäni tutkimukset on pääsääntöisesti tehty hieman eri näkökulmasta kuin tämä työ. Niissä on yleensä pyritty mallintamaan esimerkiksi sitä, tulisiko jonkun tietyn lottopelin palautusprosenttia muuttaa, jotta peli generoisi mahdollisimman suuren tulovirran pelin toimeenpanijalle.

En löytänyt tieteellisistä julkaisuista sellaisia lottopelien mallintamiseen liittyviä tutkimuksia, joiden lähtökohtana olisi ollut sama kuin tässä työssä. Syitä tähän voin vain arvailla. Yhtenä selittäjänä lienee Suomessa käytössä olevan voittorahaston hyödyntäminen Loton poteissa. Tietääkseni täysin vastaavaa järjestelmää ei ole aktiivisessa käytössä muilla lottopelejä järjestävillä toimijoilla.

4.1 LOTTOPELIEN TUTKIMUSTEN TAVOITTEISTA, MENETELMISTÄ JA TULOKSISTA

Valtaosa löytämistäni lottopelien mallintamiseen liittyvistä tutkimuksista käsitteli sitä, kannattaisiko tutkittavana olevan lottopelin palautusprosenttia muuttaa pelikat-

teen optimaalisuuden kannalta. Tämän lisäksi tutkimuskohteena on ollut se, tulisiko pelin vaikeusastetta muuttaa tai mahdollisesti lisätä peliin toinen viikoittainen arvonta. Lähtökohtaisesti näiden tutkimusten tavoite on selkeästi eronnut tämän työn tavoitteista, mutta niiden menetelmät ja tunnistetut pelaamiseen vaikuttavat tekijät ovat suuresti auttaneet tämän työn laatimisessa. Seuraavassa olen lyhyesti kuvannut kolmen lottopelijä käsittelevän tutkimuksen tavoitteita, menetelmiä sekä tuloksia. Tässä käsitellyt julkaisut ovat:

- Buying a dream: Alternative models of demand for lotto (Forrest, Simmons & Chesters; 2002)
- An Economist's Guide to Lottery Design (Walker & Young; 2001)
- What jackpot? The optimal lottery tax (Beenstock, Goldin & Haitovsky; 2000)

Useissa lottopelien myynnin mallintamista käsittelevissä tutkimuksissa lähtökohtana on ollut mallintaa loton vaihtoa siten että pääselittäjäksi on muodostettu yksittäisen pelirivin odotusarvo. Tällä tavoin malleilla on pystytty tutkimaan myös muiden, kuin päävoittoluokan vaikutusta pelin kiinnostavuuteen. Walker ja Young sovelsivat tämänkaltaista lähtökohtaa, vuonna 2001 julkaistussa tutkimuksessa: An Economist's Guide to Lottery Design (Walker & Young; 2001). Tavoitteena heillä oli selvittää, kuinka Ison-Britannian lottopeliä tulisi operoida (millainen tulisi pelin voittorakenteen, palautusprosentin ja vaikeusasteen olla), jotta sen tuotto yhteiskunnalle olisi mahdollisimman suuri. Pelirivin odotusarvon lisäksi Walker ja Young hyödynsivät pelin voittorakenteen varianssia ja vinoutta mallin selittäjinä. Voittorakenteen varianssi ja vinous ovat johdettavissa yksittäisen pelirivin odotusarvosta.

Walker ja Young rakensivat loton vaihdolle mallin, jonka avulla he pyrkivät simuloimaan sitä, tulisiko pelin edellä mainittuja parametreja muuttaa mahdollisimman kannattavan pelin luomiseksi. Mallintamisen he toteuttivat perinteisen lineaarisen regressiomallin avulla, jonka parametrit arvioitiin pienimmän neliösumman menetelmällä. Myynnin mallissa heillä oli mukana muun muassa seuraavia selittäjiä, listattuna myös mielenkiintoisimmat havainnot:

- Rivin odotusarvo - positiivinen vaikutus pelaamiseen
 - Muuttujan arvo riippui pääsääntöisesti siitä, oliko kussakin päävoitossa edellisiltä kierroksilta siirtynyttä rolloveria. Muuttuja voidaan rinnastaa potin suuruuteen
 - Positiivinen vaikutus luonnollinen, edellisiltä kierroksilta siirtyvät jättipotit perinteisesti kasvattavat myyntiä kun jättipotti kasvaa
- Voitonjaon varianssi - negatiivinen vaikutus pelaamiseen
 - Rivin odotusarvon toinen momentti
 - Riskin kasvaminen ei houkuttele pelaajia pelaamaan → negatiivinen vaikutus myyntiin
 - Selittäjän tulkinta haasteellinen

- Voitonjaon vinous - lievä positiivinen vaikutus
 - Rivin odotusarvon kolmas momentti
 - Vaikuttaisi siltä, että pelaaja hieman suosivat vinon voitonjaon tilannetta
 - Selittäjän tulkinta haasteellinen
- Myyntipisteiden lukumäärä - positiivinen vaikutus
 - Tulosta voidaan pitää luonnollisena. Mitä useammassa paikassa peliä voi pelata, sitä parempi lopputulos pelin tuoton kannalta. Vastaavaa vaikutusta ei kuitenkaan Suomesta kovin merkittävässä määrin löytynyt
- Trendi (viikkoa ensimmäisestä arvonnasta) - negatiivinen vaikutus
 - Lottopeleille suhteellisen tyypillinen tulos. Pelin kiinnostavuus laskee ajan kuluessa. Vastaavia tuloksia myös Suomessa kertapelaamisen osalta
- Vuodenaikaan sidotut I/O muuttujat
 - Vuoden viimeisellä neljänneksellä heikoin myynti. Suomessa kesä ollut perinteisesti heikointa aikaa Loton pelaamisen kannalta

Kaavassa (7) on kuvattu approksimaatio sille, kuinka Walker & Young määrittivät rivin odotusarvon yksittäisessä arvonnassa. Kaavan vasemmanpuoleinen osa kattaa pelin parimutuaaliset(muuttuvat)- ja oikeanpuoleinen osa vastaavasti kiinteät voitot:

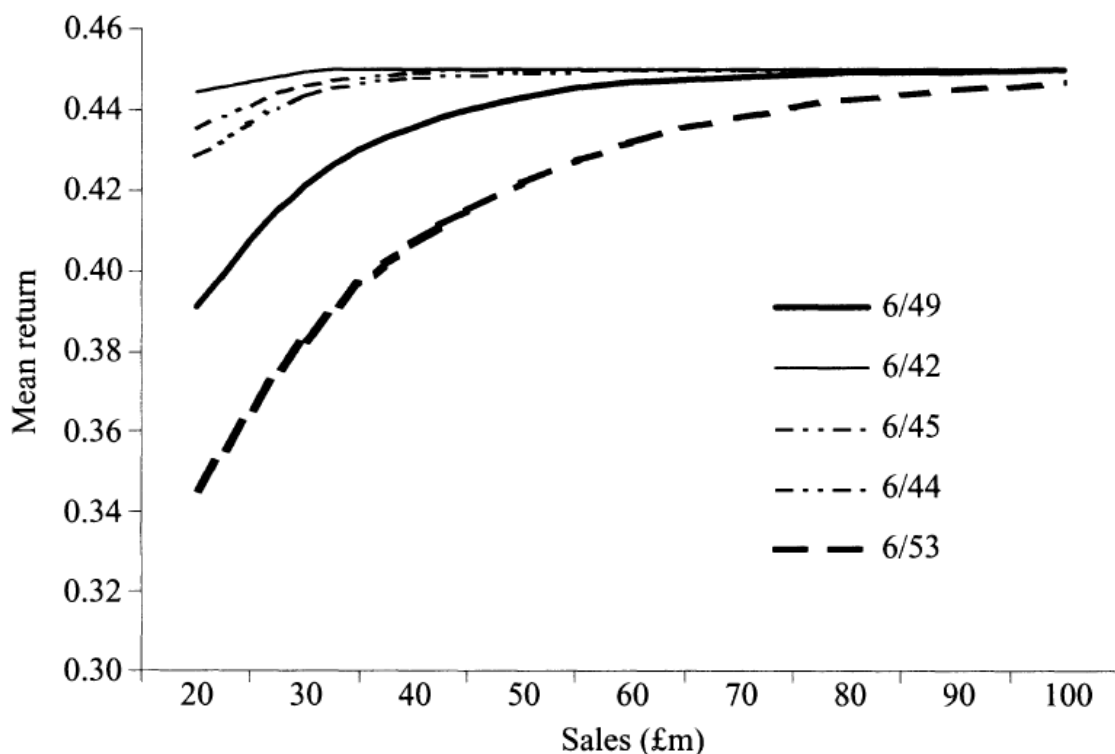
$$E(rivi) = \sum_i J_i(1 - e^{-S\pi_i})/S + \sum_i \pi_i W_i \quad (7)$$

, missä J_i kuvaa voittoluokan i jakosummaa, S arvontaan osallistuvien rivien lukumäärää, π_i voittoluokan i saavuttamisen todennäköisyyttä yksittäisellä pelirivillä ja W_i voittoluokan i kiinteän voiton suuruutta. Voitonjaon varianssi ja vinous ovat edellä kuvatun odotusarvon toiset ja kolmannet momentit.

Kaavasta (7) voidaan huomata, että yhtenä selittäjänä rivin odotusarvolle on arvontaan osallistuvien rivien lukumäärä. Tämä on juuri se, mitä loton myynnin mallilla ollaan tässä tutkimuksessa selittämässä, eli yksi mallin selittäjistä sisältää jo selitettävän suureen. Tästä johtuen tämänkaltainen matemaattinen malli ei olisi tullut kysymykseen tässä työssä. Se soveltuu hyvin malleihin, joissa loton mallinnuksen ensisijaisena tehtävä ei ole ennustaa kierroskohtaista myyntiä tai vaihtoa, vaan tarkastella esimerkiksi pelin rakenteen optimaalisuutta muun muassa palautusprosentin osalta.

Lottorivin odotusarvon (joka määriteltiin kaavassa (7)) ominaisuuksia havainnollistaa hyvin kuva 3 (Walker & Young; 2001). Siinä on havainnollistettu lottopeleille tyypillistä ja lottotutkimuksissa tiuhaan esiintyvää termiä lottopelien poikkeuksellinen mitta-kaavaetu (Lotto's peculiar economies of scale). Kuvasta nähdään, että mikäli pelissä ei ole rolloveria sisässä (kuvan oletus), niin mitä enemmän peliä pelataan, sitä suuremmaksi rivin odotusarvo muodostuu. Lopulta odotusarvo lähenee pelin laskennal-

lista palautusprosenttia, mikä kuvassa 3 on 45% (esimerkissä hyödynnetty Ison-Britannian lottoa). Lottopelien poikkeuksellinen mittakaavaetu perustuu siihen, että mitä enemmän rivejä arvontaan osallistuu, sitä pienempi on rolloverin todennäköisyys → päävoitto jaetaan pelaajille tarkasteltavalla kierroksella sitä todennäköisemmin. Tässä yhteydessä on syytä huomauttaa, että edellä mainittu poikkeuksellinen mittakaava etu ei Suomessa nykypäivänä aivan yksi yhteen pidä paikkaansa. Tämä perustuu siihen, että Veikkaus on Lotossa jo pidempään taannut pelaajille mainostamansa pottien suuruuden, eli potti ei voi olla alle mainostetun. Tästä seuraa se, että vaikka ainoastaan yksi pelirivi osallistuisi arvontaan, olisi sillä mahdollisuus voittaa markkinoitava potti. Suomessa yksittäisen pelirivin kannalta olisi siis teoriassa edullista, mikäli muita pelirivejä arvontaan osallistuisi mahdollisimman vähän. Suomen Lotossa alla kuvattu poikkeuksellinen mittakaava etu omaisi hieman toisenlaisen muodon, jossa alhaisemmilla vaihtotasolla pelaajan pelirivin odotusarvo olisikin korkeampi. Edellä kuvatunlainen muutos Loton mittakaavaetuun Suomessa on kuitenkin täysin teoreettinen. Mikäli pelin vaihtotasot laskisivat, laskettaisiin myös ajan myötä pelin potteja ja näin ollen mittakaavaetu palautuisi muistuttamaan kuvan kolme esimerkkiä.



Kuva 3: Loton poikkeuksellinen mittakaavaetu

Analyyseissään Walker ja Young tulevat tulokseen että Ison-Britannian lottopelin tuottoja voitaisiin mahdollisesti parantaa, jos pelin palautusprosenttia laskettaisiin. Tutkimuksessaan he toteavat, että tällöin pelin vaihto todennäköisesti hieman laskisi, mutta pelaajille palautettavan osuuden pienentäminen kompensoisi tämän ja jopa

kasvattaisi tuottoja. Tämän lisäksi he tulivat tulokseen, että pelin päävoiton vaikeuttaminen ei todennäköisesti kasvattaisi pelin tuottavuutta, vaikka se lisäisi jättipotti-en lukumäärää. Tätä he perustelivat sillä, että pelin vaikeutuessa rivien odotusarvo heikkenee ja rivien odotusarvon varianssi sekä vinous kasvaa. Kaksi ensimmäistä edellä mainituista vaikutuksista oletettavasti laskisivat pelin tuottoja ja ainoastaan vinouden kasvu nostaisi niitä. Walker ja Young eivät uskoneet, että yksistään vinouden kasvu olisi riittävä kompensoimaan kahden muun tekijän negatiivista vaikutusta.

Beenstock, Goldin ja Haitovsky tutkivat lottoa hyvin mielenkiintoisesta näkökulmasta julkaisussaan: *What jackpot? The optimal lottery tax* (Beenstock, Goldin & Haitovsky; 2000). He pyrkivät osoittamaan, että lottopelin optimaalisuuden kannalta ei ole järkevää, että pelin palautusprosentti on vakio kierroksesta toiseen. Tämän lisäksi he tutkivat, millä ehdoilla lottopeli kannattaisi järjestää kahdesti viikossa, perinteisen yhden arvonnän sijaan. Tutkimuskohteena heillä oli Israelissa pelattava kansallinen lottopeli.

Tutkimuksissaan Beenstock et. al. päätyvät tulokseen, että kierroksesta toiseen kiinnitetty palautus ei ole optimaalinen lottopelissä, vaan paras tuotto saavutettaisiin, jos pelin palautus vaihtelisi pelin vaiheen mukaan (mikä on pelin pottityyppi). Heidän tuloksiensa perusteella lottopeleissä tulisi pyrkiä siihen, että ensimmäisillä kierroksilla rolloverin todennäköisyys olisi mahdollisimman suuri, eli vaihto alhainen. Tähän pyritään rajoittamalla ensimmäisillä kierroksilla jättipotin suuruutta, pitämällä se kuitenkin niin suurena, että näiltä kierroksilta siirtyvä rollover on pelaajille merkittävä tulevilla kierroksilla.

Beenstock et. al. olettavat Israelin dataan vedoten, että kun riittävän monta jättipottia on jäänyt jakamatta pelaajiin iskee "lottomania". Lottomaniasta puhutaan myös esimerkiksi edellä esitetyissä Walkerin ja Youngin tutkimuksissa. Tällöin pelaaminen kasvaa epälineaarisesti ja tuolloin kannattaisi pelin palautusta nostaa. Tässä tutkimuksessa on tarkasteltu palautusta ainoastaan päävoittoluokan osalta. Tutkimusmenetelmänä oli simulointi, jossa loton rollovereiden lukumäärä oli rajoitettu kolmeen. Taulukossa 5 on kuvattu ote tutkimuksen simulaatioiden tuloksista:

Summary of simulations

	Games per week	Option	α	Expected weekly profit (shekels millions)
1	1	F	0.21	5.218
2	1	O	0.25	5.275
3	1	F	0.32	5.060
4	1	E	0.23	5.383
5	1	G	varies	5.538

Taulukko 5: Loton optimaalisesta palautuksesta (Beestock et. al.; 2000)

Taulukossa α kuvaa päävoittoluokan osuutta pelin vaihdosta. Alemmille voittoluokille palautettava osuus oli kiinteä 34% $\rightarrow$ pelin palautusprosentti $\alpha_k = \alpha + 34\%$.

Taulukon ensimmäinen rivi kuvaa simulaatiota, joka on toteutettu Israelin tutkimuksen suorittamisen aikana voimassa olleiden peliehtojen mukaan: päävoittoluokan osuus vaihdosta on ollut kierroksesta toiseen kiinnitetty 21%. Simulointien mukaan odotettavissa oleva viikoittainen tuotto on ollut noin 5,2 miljoonaa shekeliä. Simuloinnissa 2 optimoitiin päävoittoluokan osuutta vaihdosta. Optimaaliseksi päävoittoluokan osuudeksi saatiin tässä simuloinnissa 25%. Simulointi 2 toteutettiin seuraavalla rajoitusehdolla: päävoittoluokan osuuden tulee olla kiinteä kierroksesta toiseen, eli kyseessä oli edelleen perinteinen lottopeli. Odotettavissa oleva tuotto kasvoi noin prosentin simulointiin 1 nähden, eli merkittävää hyötyä ei päävoittoluokan osuuden kasvattamisesta saatu. Simulointi 3 toimi esimerkkinä siitä, että mallin odotettavissa olevat tuotot laskevat kun päävoittoluokan osuus kasvaa liian korkeaksi, sillä samalla pelin palautusprosentti nousee. Simuloinnissa 4 jättipottien suuruudet oli asetettu kiinteiksi ja ne oli ennalta määrätty. Tällainen lähtökohta on tietystä mielessä rinnastettavissa tämän työn menetelmiin. Lopputuloksena neljännessä simuloinnissa oli, että kiinteiden päävoittojen mallissa odotettavissa oleva tuotto kasvaa noin 3%, mitä voidaan pitää jo suhteellisen merkittävänä tuloksena. Simulointien mukaan parhaat tuotot saavutettaisiin, mikäli pelin palautusta vaihdeltaisiin voimakkaasti eri jättipottikierrosten välillä. Optimaaliseksi jättipottirakenteeksi muodostui: $J_0 = 4,3M$; $J_1 = 4,0M$; $J_2 = 7,2M$; $J_3 = 26,4M$, missä alaindeksi kuvaa jättipotissa olevien rolloverien lukumäärää. Erittäin mielenkiintoisena seikkana nähdään, että toisen viikon jättipotti on pienempi kuin ensimmäisen viikon jättipotti $4,0M < 4,3M$. Tulos on yllättävä, sillä toisen viikon jättipotti sisältää myös ensimmäisellä viikolla päävoittoon siirretyt varat. Tämä selittyy sillä, että tässä simulaatiossa kunkin viikon jättipotti J_i määritettiin kaavalla $J_i = \alpha_i P \sum_i R_i$, missä α_i kuvaa päävoittoluokan palautusprosenttia kun jättipotissa on mukana i rolloveria, P kuvaa rivihintaa ja R_i kierroksen i arvontaan osallistuvaa kokonaisrivimäärää. Eli näin ollen jos α_1 on esimerkiksi pienempi kuin $\frac{1}{2} \alpha_0$, niin on hyvin mahdollista, että $J_1 < J_0$. Mielestäni tämä käy kuitenkin hyvästä esimerkistä siitä, että vaikka simulaatioiden valossa näyttäisi, että tällainen malli voisi olla järkevä, ei näin kuitenkaan todellisuudessa ole. Ihmisten reaktio siihen, että jättipotti lottopelissä pienenee, vaikka voittajia ei löydy, olisi mitä todennäköisimmin niin negatiivinen, että peli joutuisi todella suuriin vaikeuksiin. Tällaista jättipottirakennetta ei tietääkseni ole kuitenkaan missään toimeenpantavassa lottopelissä sovellettu ja näin ollen sen todellinen toimivuus jää ainoastaan arvailujen varaan.

Tutkimuksissaan Beenstock et. al. pyrkivät myös selvittämään, millä ehdoin kannattaisi kerran viikossa arvottavasta lottopelistä tehdä kahdesti viikossa arvottava. Tuloksia ei voida heidän omien sanojensa mukaan pitää täysin luotettavina, koska ne perustuvat yhden arvonnin pelin mallinnostuloksiin. Ajatusmalleina tulokset ovat kuitenkin hyvin kiinnostavia. Beestock et. al. esittävät että viikkotason myynnin tulisi

kasvaa kertaluokkaa 30%, jotta muutos kahteen arvontaan olisi kannattava. Tämän perusteena oli lähinnä se, että vaikka viikkotason myynti toisen arvonnän myötä nousisi, arvonnasta toiseen siirtyvien jättipottien taso laskisi → vähemmän lotto-
maniaa ja kierroksia, jolloin peliä pelattaisiin todella paljon. Edellä kuvattu pätee lähinnä lottoputkien alkupäähän, eli jättipotit kasvaisivat alussa hitaammin. Mutta koska ne kasvaisivat hitaammin, niin tällöin myös rollovereiden todennäköisyys kasvaisi, mikä kompensoisi osaltaan pienentyneitä jättipotteja. Keskimäärin pelissä voitettun jättipotin suuruus ei kuitenkaan muuttuisi, vaikka arvontakohtaisen vaihdon taso laskisi (ks. Kaava (4)). Maailmalla on onnistuneita esimerkkejä siitä, että lotossa on otettu käyttöön toinen viikoittainen arvonta. Yhdistyneen Kuningaskunnan lottopelissä otettiin toinen arvonta vuonna 1997 ja tämän johdosta pelin myynti kasvoi noin 30%. Tämän lisäksi muun muassa monikansallisessa lottopelien logiikan omaavassa Euromillions-pelissä on toisella viikoittaisella arvonnalla saatu aikaan hyvinkin positiivisia tuloksia.

Käsittääkseni Beenstockin et. al. kahden arvonnän simulaatioissa oli ajateltu erottaa arvontojen rahankierto toisistaan (eli käytännössä luotiin kaksi eri peliä). Tämä tarkoittaa sitä, että mahdolliset rolloverit eivät siirtyisi seuraavaan arvontaan, vaan seuraavaan samana päivänä arvottavaan arvontaan. Kirjoittajan mielestä tällainen tapa toteuttaa toinen lottoarvonta ei kuitenkaan välttämättä ole se järkevin tapa. Enemmän kannattaisi ehkä pyrkiä siihen, että jättipotit siirtyvät arvonnasta toiseen ja yksi peli omaisi ainoastaan kaksi viikoittaista arvontaa. Molemmilla tavoilla varmasti on puolensa, eikä tähän problematiikkaan ole tässä tarkoitus sen syvemmin paneutua.

Forrest, Simmons ja Chesters (Forrest et. al.; 2002) lähestyvät Loton myynnin mallintamista hieman eri näkökulmasta kuin esimerkiksi Walker & Young. Heidän hypoteesinaan on, että pelaajien motiivi pelaamiseen on enneminkin mahdollisuus haaveilla suurista voitoista ja siitä, kuinka he mahdollisesti tällaisen voiton käyttäisivät. He myös totesivat, että pelaamiskäyttäytymisen mallintaminen rivin odotusarvo pääselittäjänä perustuu tuoton odotusarvo teoriaan ("expected utility theory"), mikä on ristiriidassa sen kanssa että lottopelejä pelatessaan ihmiset tietoisesti osallistuvat negatiivisen odotusarvon peliin, mutta saattavat kuitenkin samanaikaisesti vakuuttaa esimerkiksi kotinsa. Ongelmaksi muodostuu siis: miksi riskiä karttavat ihmiset (vakuutuksenottajat) osallistuvat ylipäättään lottopeleihin? Tähänkin kysymykseen on varmasti olemassa taloustieteellisiä näkökulmia, Forrest, Simmons ja Chesters halusivat kuitenkin mallintaa pelaajien käyttäytymistä selittäjillä, jotka heidän näkemyksensä mukaan parhaiten kuvaisivat pelaajien todellisia motiiveja.

Forrest, Simmons ja Chesters lähestyivät lottopelien pelaamismotiiveja näkökulmasta, että ne antavat mahdollisuuden unelmoida jostain todella suuresta, mitä heidän olisi muuten vaikea saavuttaa (kirjoittajan huomautus: lähes jokaisella ihmisellä luultavasti olisi paremmat mahdollisuudet saada suuri omaisuus jollain muulla tavalla

kuin lottoamalla. Muut tavat vaan ovat pääsääntöisesti huomattavasti työläämpiä kuin muutaman lottorivin pelaaminen.). Veikkauksen asiakkailleen teettämässä kyse-lyssä mahdollisuus haaveilla suuresta voitosta on järjestään ollut merkittävin motiivi Loton pelaamisessa. Näin ollen tätä työtä tehdessäni, koin Forrestin et. al. lähesty-mistavan Loton mallintamiseen kaikista antoisimmaksi.

Tutkimuksessaan Forrest et. al. hyödynsivät Yhdistyneiden Kuningaskuntien lottopelin dataa. Loton myynnin mallin pääselittäjäksi he valitsivat niin jättipotin koon kuin yksittäisen rivin odotusarvonkin. Edellä mainittujen pääselittäjien varaan Forrest et. al. rakensivat kaksi kilpailevaa mallinnosta lottopelin myynnille ja vertailivat niitä keskenään, toisessa mallissa pääselittäjänä oli jättipotin koko ja toisessa taas yksit-täisen pelirivin odotusarvo. Matemaattiset mallit olivat kaksivaiheisia pienimmän ne-liösumman malleja (two-stage least squares) ja toteutettu erikseen sekä keskiviikko-että lauantaiarvonnoille. Seuraavassa on kuvattu lauantaiarvonnan malli jättipotti pääselittäjänä:

Vaihe 1

$$J = f(vakio, q_{t-1}^l, Q_k, TREND, TREND^2, J_{ex}, DIANA, ROLL) \quad (8)$$

Vaihe 2

$$q = f(vakio, q_{t-1}^l, Q_k, TREND, TREND^2, J_{ex}, DIANA, J) \quad (9)$$

, missä q kuvaa tarkasteltavan kierroksen myyntiä (luonnollinen logaritmi), J esti-maattia tarkastelukierroksen jättipotista, Q_k edellisen keskiviikon lottokierroksen myyntiä, $TREND$ myynnin muutosta ajassa, aivan kuten $TREND^2$, J_{ex} kuvaa ylimää-räistä jättipottirahaa mitä peliin on mahdollisesti lisätty, $DIANA$ on I/O muuttuja joka kuvaa kierrosta jolloin arvontapäivää siirrettiin päivällä eteenpäin prinsessa Dianan hautajaisten takia ja $ROLL$ kuvaa mahdollista rolloveria.

Tuloksissaan Forrest et. al. päätyivät seuraaviin johtopäätöksiin: Jättipotin käyttämi-nen pääselittäjänä tuottaa parempia tuloksia selitysasteen valossa kuin pelirivien odotusarvon hyödyntäminen. He päätyivät myös tulokseen, että saattaisi olla kan-nattavaa laskea Englannin lottopelin palautusprosenttia ainakin lauantai arvannon osalta. Koska jättipotin merkityksen pelin myynnille havaittiin olevan todella suuri, nähtiin tarpeelliseksi pyrkiä säilyttämään jättipottien taso mahdollisimman korkeana vaikka pelin palautusta laskettaisiinkin. Myös pelin vaikeuttamisen selvittämistä suo-siteltiin, koska tämä lisäisi jättipottien lukumäärää. Pelin vaikeuttamiseen liittyen Walker & Young esittivät edellä kuvatuissa omissa tutkimuksissaan, että Ison-Britannian loton tuotot eivät kasvaisi mikäli peliä vaikeutettaisiin (Walker & Young; 2001).

4.2 AJATUKSIA LOTTOPELIEN TUTKIMUKSISTA

Lottopelien mallintaminen julkaistuissa tutkimuksissa on lähestynyt lottopelien problematiikkaa täysin eri näkökulmasta kuin olen tässä työssä tehnyt. Näkökulma on pääsääntöisesti ollut se, voitaisiinko pelin tuottoja kasvattaa muuttamalla esimerkiksi pelin palautusprosenttia tai vaikeusastetta. Edellä mainitut muutokset lottopeliin edellyttävät pelin sääntöjen ja mahdollisesti tietoteknisen pelijärjestelmän muokkaamista ja ovat siinä mielessä kohtalaisen suuria operaatioita. Edellä lyhyesti esittämäni tulokset esitetään pääsääntöisesti hyvin luotettavina, vaikka ainakin itse suhtaudun suhteellisen skeptisesti kaikkiin matemaattisesti laadittuihin suosituksiin siitä, kuinka lottopelin peliformaattia tulisi olemassa olevaan peliin nähden muuttaa. Tämä saattaa kuulostaa oudolta, olenhan itsekkin tässä työssä lottopelien myyntiä pyrkinyt mallintamaan, mutta alla on joitain pointteja jotka mielestäni jäivät suhteellisen vähälle huomiolle tutustumissani tutkimuksissa.

Ihmisten reagointi lottopelien muutoksiin voidaan olettaa olevan enemmän tunnepitoista kuin rationaalista. Perustuuhan niiden pelaaminenkin enemmän haaveiluun siitä, mitä kaikkea ihminen voisi suuren voiton avulla tehdä, kuin mihinkään rationaaliseen sijoituspäätökseen.

Esimerkiksi peliformaatin muuttaminen siten, että päävoitosta tulee vaikeampi saavuttaa, on mielestäni erittäin haasteellista matemaattisesti mallintaa. Matemaattiset mallit joudutaan rakentamaan olemassa olevan pelin tilastojen nojalla ja näin ollen niihin ei pääsääntöisesti kyetä sisällyttämään dataa uudesta formaatista. Vastaava tilanne esiintyy tietenkin aina, kun pelien parametreissa siirrytään uudelle alueelle. Peliformaatin muutosta voidaan pitää pelaajien kannalta kuitenkin hyvin suurena muutoksena, verrattaessa esimerkiksi hinnankorotukseen tai siihen, että päävoiton osuutta voitonjaossa kasvatetaan. On tilanteita, joissa lottopelien vaikeuttaminen on perusteltua (esimerkiksi voimakkaasti kasvaneet pelaajamäärät), mutta tällaisessakin tilanteessa koen, että matemaattisilla malleilla on hyvin hankala estimoida pelaajakäyttäytymisen muutosta. Kasvaneet potit eivät välttämättä riitä kompensoimaan sitä, että pelaajat kokevat päävoiton saamisen vaikeutumisen hyvin negatiivisesti.

Walker ja Young (Walker & Young; 2001) ehdottavat edellä esitetyssä julkaisussaan, että heidän tutkimassaan lottopelissä kannattaisi harkita jopa palautusprosentin laskua. He esittävät, että vaikka pelin suosio todennäköisesti hieman laskisi, kasvaneet tuotot peliriviä kohden riittäisivät kompensoimaan tämän. Itse koen tällaisen mallinnoksen äärimmäisen haasteelliseksi. Vaikka matemaattisesti voitaisiin osoittaa, että nykydataan perustuen asia olisi näin, malleihin ei voida millään sisällyttää sitä negatiivisen julkisuuden vaikutusta, joka palautuksen laskemisesta todennäköisesti seuraisi. En tiedä, onko maailmassa toteutettu menestyneesti yhtään palautusprosentin laskemista, mutta epäilen, että niitä ei ole hirveästi edes harkittu edellä kuvatun negatiivisen julkisuuden riskin johdosta. Palautusprosenttien nostoja on useissakin pe-

leissä tehty, mutta niitä en tässä sen tarkemmin käsittele. Edellä kuvatuista syistä koenkin, että jos jonkin pelin mallinnokset selkeästi suosittelevat alentamaan pelin palautusprosenttia, kannattaa tämä mahdollisesti huomioida uusien lottopelien suunnittelussa, mutta olemassa oleviin peleihin en tällaista menetelmää käyttäisi.

Tutkimuksissa joihin perehdyin, ei juurikaan kiinnitetty huomiota siihen, että useimmiten tutkittavien parametrimuutosten myötä siirrytään alueelle, josta ei ole lainkaan tilastollista dataa saatavilla. Jos esimerkiksi päävoiton osuutta merkittävästi lottopelissä nostetaan, vaikuttaa se selkeästi esimerkiksi pelaajien kokemaan odotusarvoon, kun potti on kasvanut suureksi. Voisi olettaa, että tällaisissa tilanteissa mallinnosten luotettavuus selkeästi heikkenee. Tässä työssä ei esimerkiksi pyritty tutkimaan sitä, kuinka Loton päävoiton palautusprosentin muuttaminen pelaamiseen vaikuttaisi, koska tällainen tutkimus koettiin hyvin haasteelliseksi.

Loton mallintamista ja optimoimista käsitelleissä tutkimuksissa oli monia hyödyllisiä näkökulmia tätä työtä varten. Mutta kuten edellä kuvasin, koen että matemaattisten mallien perusteella ei yksin voida luotettavasti ennustaa peliparametristen muutosten (pelin palautusprosentti, arvontojen lukumäärä tai pelin vaikeusaste) vaikutusta pelaamiseen. Niiden tueksi tarvitaan muutakin informaatiota, kuten esimerkiksi kyselytutkimuksia tai fokusryhmiä ja toisaalta myös aimo annos kokeneiden asiantuntijoiden näkemystä. Ja vaikka kaikkia edellä mainittuja menetelmiä hyödynnettäisiin, ei silti voitaisi olla varmoja siitä, että pelin muuttaminen lisäisi pelaajien mielenkiintoa sitä kohtaan.

5 LOTON MYYNNIN MALLINTAMISEN MENETELMÄ JA SIMULAATTORIN TOIMINTAPERIAATE

Tässä diplomityössä Loton myyntiä on mallinnettu usean selittäjän regressiomallien avulla. Päätöstä tuki asiantuntija-arvioiden lisäksi aiheesta julkaistut tieteelliset tutkimukset, joissa hyödynnettiin pääsääntöisesti vastaavia menetelmiä lottopelien myynnin mallintamisessa.

Harkinnassa oli myös aikasarjamallien hyödyntäminen, mutta koska alustavat tulokset eivät indikoineet paremmasta mallinnoskyvystä, päätettiin myynnin mallintaminen ja simulaattori rakentaa usean selittäjän regressiomallien varaan. Neuroverkko-malleja ei kovin pitkälle edes harkittu, koska niiden avulla ei selittäjien vaikutuksesta myyntiin olisi saatu oikein mitään irti.

Usean selittäjän regressiomalleja koskeva teoria on koostettu teoksen (Pindyck, Rubinfeld; 1991) ja Ilkka Mellinin luontomateriaalien ”Lineaarinen regressioanalyysi” & ”Yleinen lineaarinen malli” perusteella. Asiasta syvällisemmin kiinnostunut lukija ohjataankin tutustumaan tarkemmin edellä mainittuihin materiaaleihin.

5.1 USEAN SELITTÄJÄN REGRESSIOMALLIN RAKENNE, OLETUKSET JA PARAMETRIEN ESTIMOINTI

Tässä osiossa esitetyt oletukset perustuvat sekä Mellinin luentomateriaaleihin ”Lineaarinen regressioanalyysi” & ”Yleinen lineaarinen malli”, että teokseen *Econometric Models & Economic Forecasts* (Pindyck, Rubinfeld; 1991).

Usean selittäjän regressiomalli voidaan kuvata kaavassa (10) esitetyllä tavalla.

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i \quad (10)$$

, missä y_i kuvaa riippuvan muuttujan havaintoa i , x_i selittävän muuttujan havaintoa, ε_i havainnon i virhetermiä. β_0 kuvaa vakiotermiä ja $\beta_j; j > 0$ kuvaa selittäjän x_j ei-satunnaista ja tuntematonta regressiokerrointa.

Muuttujille Y & X pätee seuraavat perusoletukset:

- Selitettävä muuttuja Y on suhdeasteikollinen satunnaismuuttuja
- Selittävät muuttujat X ovat kiinteitä eli ei-satunnaisia muuttujia

Usean selittäjän regressiomallin perusoletukset:

1. Selittäjän X_j havaitut arvot X_{ji} ovat kiinteitä eli ei satunnaisia vakioita $\forall j, i$
2. Selittäjien välillä ei esiinny lineaarisia riippuvuuksia
3. $E(\varepsilon_i) = 0 \forall i$ - mallin rakenneosassa ei ole tehty systemaattista virhettä
4. $Var(\varepsilon_i) = \sigma^2 \forall i$ - homoskedastisuusoletus, eli kaikilla virhetermeillä on sama varianssi
5. $Cor(\varepsilon_i, \varepsilon_l) = 0; i \neq l$ - virhetermit eivät korreloi keskenään
6. $\varepsilon_i \sim N(0, \sigma^2) \forall i$ - Virhetermit ovat normaalijakautuneita

Edellä esitettyjä vaatimuksia testataan muodostettujen regressiomallien tilastollisen testaamisen yhteydessä.

Kaavassa (10) esitetty usean selittäjän regressiomalli voidaan esittää matriisimuodossa seuraavasti:

$$y = X\beta + \varepsilon \quad (11)$$

Jos usean selittäjän regressiomallin perusoletusten kohta 2. on voimassa, voidaan vektorin β pienimmän neliösumman estimaatti b (PNS-estimaatti) esittää muodossa:

$$b = (X'X)^{-1}X'y \quad (12)$$

Gaussin ja Markovin lauseen mukaan edellä kuvattu pienimmän neliösumman estimaattori b on paras vektorin β lineaaristen ja harhattomien estimaattoreiden joukossa, mikäli perusoletukset 1.-5. ovat voimassa.

5.2 USEAN SELITTÄJÄN REGRESSIOMALLIN SELITTÄJIEN IDENTIFIOINTI, MALLIN DIAGNOSTINEN TESTAUS JA SILLÄ ENNUSTAMINEN

Edellisessä kappaleessa kuvattiin käytettävän regressiomallin rakenne, siihen liittyviä perusoletuksia sekä mallin parametrien estimointi. Tässä kappaleessa on lyhyesti kuvattu teoria ja toteutus siitä, kuinka mallin selittäjät on identifioitu. Tämän lisäksi osiossa on lyhyesti käsitelty regressiomallien tilastollisen testauksen teoriaa.

Kaavassa (11) esitetyn regressiomallin rakenneosa voidaan määrittää (Mellin; Lineaarinen regressioanalyysi):

$$y = X\beta \quad (13)$$

Ja edelleen voidaan asettaa regressiomallin tavoite:

$$E(y|X) = X\beta \quad (14)$$

Jotta kaavassa (14) esitetystä tavoitteesta onnistuttaisiin mahdollisimman hyvin, tulee mallin rakenneosan olla oikein spesifioitu. Toisin sanoen mallissa tulee olla mukana mahdollisimman kattavasti selitettävään muuttuun vaikuttavat tekijät. Tämän lisäksi mallissa ei saisi olla turhia selittäjiä.

- Jos mallista puuttuu olennaisia selittäjiä → regressiokertoimien PNS-estimaattoreista tulee yleensä harhaisia
- Turhat selittäjät aiheuttavat tehottomuutta PNS-estimaattoreihin → kertoimien variansseista tulee tarpeettoman suuria

Estimaattoreiden harhaisuutta voidaan pitää haitallisempana kuin tehottomuutta.

Regressiomallien identifioinnissa voidaan hyödyntää erilaisia mallivalintastrategioita. Tässä työssä hyödynsin menetelmää, jota voidaan kutsua alaspäin askellukseksi. Tämän prosessin alussa identifioitiin mahdollisimman kattavasti potentiaaliset selittäjäkandidaatit. Tämä toteutettiin hyödyntämällä edellä kuvattuja kansainvälisiä tutkimuksia, työn laatijan omaa näkemystä, sekä Veikkaus Oy:n henkilökunnan asiantuntija-arvioita. Alaspäin askelluksen menetelmä voidaan kuvata seuraavasti:

1. Lähtömalliin sisällytetään kaikki selittäjäkandidaatit
2. Valitaan mallinvalintatesteissä käytettävä merkitsevyystaso *out*
3. Estimoidaan malli niiden selittäjien avulla, jotka mallissa ovat
4. Testataan löytyykö selittäjiä merkitsevyystasoa *out* vasten, jotka mallista tulisi poistaa
5. Jos poistettavia selittäjiä ei löydy → malli valmis
6. Poistetaan mallista selittäjä, jonka p-arvo on suurin
7. Palataan vaiheeseen 3

Tässä työssä hyödynnettiin pääsääntöisesti yllä kuvattua mallinrakennusalgoritmia. Tämän lisäksi pyrittiin hyödyntämään asiantuntija osaamista mahdollisimman paljon, jotta mallista ei tulisi ainoastaan matemaattisesti validi, vaan se olisi myös asialoogisesti tulkittavissa. Tätä pyrittiin varmistamaan sillä, että jokaiselle mukana olevalle selittäjälle kyettiin antamaan suhteellisen järkevä selitys, selittäjien parametrien suuruusluokan katsottiin olevan järkevä ja ylipäättään katsottiin kuvaako kukin parametri todellisuutta mielekkäästi.

Mallin diagnostinen testaus suoritettiin Mellinin materiaalissa Lineaarinen regressioanalyysi kuvatulla tavalla (Mellin; Lineaarinen regressioanalyysi). Seuraavassa on lyhyesti käyty läpi se mitä testattiin ja miksi:

Mallin suorituskykyä mittaavalla diagnostiikalla pyrittiin ennen kaikkea testaamaan kuinka hyvin edellä kuvatut regressiomallien perusoletukset 1.-6. ovat voimassa. Mallille suoritettiin seuraavia diagnostisia testauksia:

1. Graafinen tarkastelu regressiomallien mallien estimointikyvystä
2. Poikkeavien havaintojen etsintä ja regressiokertoimien vakioisuus
3. Selittäjien multikollineaarisuus
4. Residuaalien normaalius, homoskedastisuus ja autokorrelaatio
5. Mallien ennustuskyky

Muodostetuilla regressiomalleilla voidaan ennustaa Loton myyntiä, kun mallien selittäjien parametrit ovat tiedossa. Mallien tuottamaksi parhaaksi ennusteeksi voidaan osoittaa muodostuvan (Mellin; Lineaarinen regressioanalyysi)

$$y_e = \beta_0 + \beta_1 x_{1h} + \beta_2 x_{2h} + \dots + \beta_k x_{kh} \quad (15)$$

, missä y_e kuvaa selitettävän suureen ennustetta, ja $x_{jh}; j = 1 \dots k$, kuvaa selittäjien havaittuja arvoja. Edellä kuvattujen ennusteiden avulla voidaan edelleen estimoida kunkin kierroksen eri kestojen myyntien estimaatit ja tämän jälkeen määrittää kierroksen vaihdolle estimaatti kaavan (6) avulla.

5.3 SIMULAATTORIN TAVOITTEET JA TOIMINTAPERIAATE

Tämän työn eräänä keskeisimpänä tavoitteena oli muodostaa simulaattori, jonka avulla pystyttäisiin simuloimaan Loton vuositasen (52 kierrosta) myyntejä. Vuositason myyntien simulointien ohessa on tarkoitus tutkia myös sitä, kuinka voittorahaston kulutuksen vuotuiset odotusarvot vaihtelevat erilaisilla pottitarjoamilla. Simulaattorin avulla tullaan tutkimaan muun muassa sitä, miten erilaisten Loton pottitarjoamien vuositasen myyntien ja voittorahaston käytön odotusarvot poikkeavat toisistaan. Tämän lisäksi tarkastelun alla on se, miten Loton myynnin ja voittorahaston käytön jakaumat poikkeavat toisistaan eri pottitarjoamilla.

Simulaattorin toimintaperiaatteena on Monte Carlo simulointi, jossa simuloidaan yksittäisten lottoarvontojen tuloksia päävoiton löytymisen osalta ja edelleen tämän perusteella vuositason myyntiä ja voittorahaston kulutusta. Simulaattorin tärkeimpänä tutkimuskohteena on ennalta määritettyjen pottitarjoamien vertailu niin vuositason myynnin kuin voittorahaston kulutuksenkin osalta. Tässä työssä tutkimuksen kohteena olleet pottitarjoamat on esitetty taulukossa 14. Simulaattorin lähtökohtana käytettiin vuoden 2009 10 viimeisen viikon loton tunnuslukuja (potti, pottityyppi jne...), joten näin ollen voidaan ajatella, että simuloinnin kohteena oli vuoden 2010 Loton myynti ja voittorahaston kulutus. Simulaattorin toiminta on esitetty yksityiskohtaisemmin kappaleessa 8.1.

6 MALLINNOKSISSA KÄYTETYN AINEISTON KUVAUS

Tässä osiossa on kuvailtu käytössä ollutta aineistoa, jonka pohjalta usean selittäjän regressiomallit lopulta identifioitiin. Tarkastelun kohteena on otos mielenkiintoisimmista selittäjistä, selitettävän muuttujan lisäksi.

Regressioanalyysissä käytetty data kerättiin pääsääntöisesti Veikkauksen tilastollisista tietopankeista. Tämän lisäksi hyödynnettiin soveltuvien osien myös tilastokeskuksen tarjoamia tietokantoja. Analyysit suoritettiin ajanjaksona 11/2002-53/2009 kerätystä aineistosta. Havaintoja aineistossa oli yhteensä 408 kappaletta. Verrokkiaineistona käytettiin kierrosten 1/2010-52/2010 myyntidataa. Verrokkiaineistoa käytettiin muun muassa muodostettujen regressiomallien selityskyvyn testaamiseen.

6.1 TILASTOLLISEN AINEISTO JA SEN TULKINTAA

Regressiomallien selittäjäkandidaatit identifioitiin niin Veikkauksen työntekijöiden asiantuntija-arvioiden, aiheeseen liittyvien tieteellisten julkaisujen kuin työn tekijän omien ajatustenkin pohjalta. Selitettävänä muuttujina tässä työssä olivat Loton myynnit kestoittain. Taulukosta 6 voidaan nähdä analyysidatan tilastolliset tunnusluvut 1,2,3,5 ja 10 viikon kestopelien myynneille. Taulukossa 6 on myös esitetty kahden selittäjäkandidaatin: potti ja myyntipaikkojen lukumäärä tunnuslukuja. Taulukossa 7 on esitetty ei-suhdeasteikollisten selittäjien tilastollisia tunnuslukuja. Kuvasa 4 on havainnollistettu selitettävien muuttujien ja tärkeimmän selittäjän eli potin jakaumia histogrammien avulla.

Taulukon 6 ja kuvan 4 histogrammin avulla voidaan Loton kestojen myynneistä tehdä seuraavia havaintoja:

- Yhden viikon pelaaminen on tarkasteluajanjaksona ollut euromääräisesti selkeästi suosituinta. Kertapelaamisessa on havaittavissa myös joitain niin sanottua "lottomanian" piirteitä: jopa 15 miljoonan euron kierrosmyyntejä
- 2, 3 ja 5 viikon kestopelaaminen on kaikista stabiileinta. Kahden ja kolmen viikon kestopelaaminen on silmämääräisesti suhteellisen normaalijakautunutta

- 10 viikon kestopelaaminen omaa hyvin vahvan perustason. Siinä on ajoittain myös selkeitä piikkejä, esimerkiksi hinnankorotuksista johtuen
- Yksikään jakauma ei ole negatiivisesti vino, eritoten yhden ja kymmenen viikon jakaumat ovat positiivisesti vinoja
 - Tulosta voidaan pitää odotettuna. Kaikilla kestoilla on suhteellisen vakaa minimitaso, suuren potin tai esim. hinnankorotuksen yhteydessä selkeää positiivista vaihtelua → positiivinen vinous
- Kaikkien kestojen jakaumat ovat positiivisesti huipukkaita, eritoten yhden ja kymmenen viikon kestojen myynnit

	M_M€_1vko	M_M€_2vko	M_M€_3vko	M_M€_5vko	M_M€_10vko	Potti_M€	Myyntip_1000
Keskiarvo	6,08	0,25	0,18	0,60	1,11	1,68	3,26
Keskihajonta	1,38	0,03	0,02	0,08	0,57	1,20	0,15
Minimi	4,53	0,19	0,13	0,37	0,30	0,65	2,89
Maksimi	15,52	0,37	0,26	0,90	6,23	6,90	3,60
10% alle	4,89	0,22	0,15	0,50	0,64	0,70	3,09
25% alle	5,25	0,23	0,16	0,55	0,81	0,70	3,13
50% alle	5,82	0,25	0,18	0,60	0,97	1,30	3,28
75% alle	6,46	0,27	0,19	0,65	1,29	2,20	3,37
90% alle	7,36	0,29	0,21	0,68	1,64	3,33	3,45
Vinous	3,17	0,61	0,73	0,19	3,80	1,50	-0,22
huipukkuus	15,79	0,94	0,83	0,97	24,47	2,10	-0,59

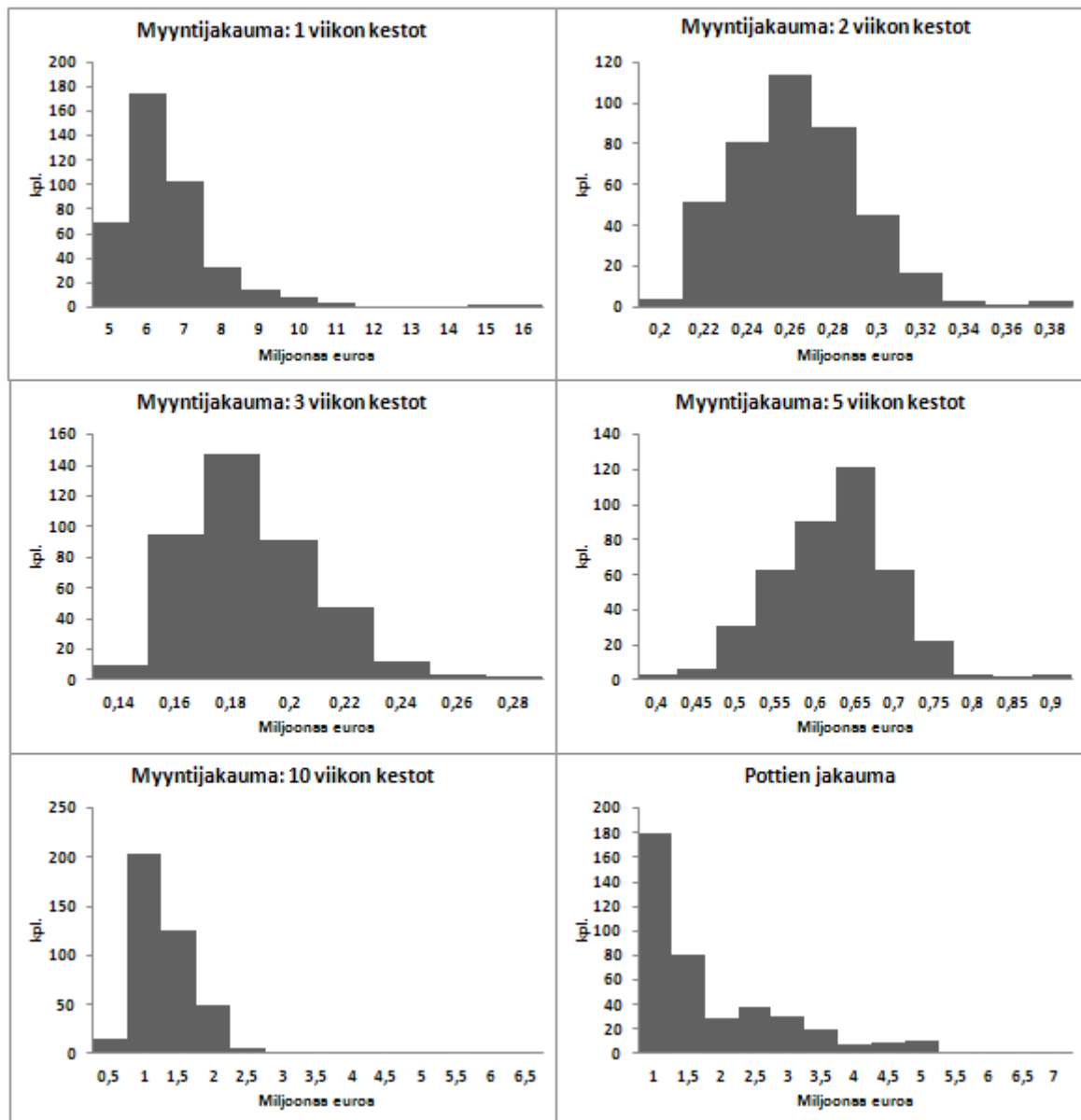
Taulukko 6: Tilastollisen aineiston tunnuslukuja (suhdeasteikolliset muuttujat)

pottityyppi	kpl.	osuus-%	hinta	kpl.	I/O	hinnankor.	ylim. potti	arvonta sun.
0	214	52,5 %	0,60	100	0	406	372	377
1	106	26,0 %	0,70	208	1	2	36	31
2	56	13,7 %	0,80	100				
3	23	5,6 %						
4	7	1,7 %						
5	2	0,5 %						
Yht	408	100,00 %		408		408	408	408

Taulukko 7: Tilastollisen aineiston tunnuslukuja (ei suhdeasteikolliset muuttujat)

Taulukoiden 6 ja 7 sekä kuvan 4 histogrammin avulla voidaan tässä käsitellyistä se-littäjistä tehdä seuraavia havaintoja:

- Potti on tarkasteluajanjaksolla vaihdellut hyvinkin laajalti: 0,65M€-6,90M€.
- Valtaosa poteista on ollut alle 4M€
- Myyntipaikkojen vaihteluväli tarkasteluajanjaksolla ollut luokkaa: 2900-3600
- Yli puolet poteista (52,5%) ollut niin sanottuja peruspotteja
- Tarkasteluajanjaksona on ollut kaksi hinnankorotusta
- Ylimääräinen potti on arvottu 36 kertaa
- Loton arvonta on suoritettu sunnuntaisin 31 kertaa



Kuva 4: Tilastollisen aineiston graafinen esitys

7 LOTON MYYNIN MALLINTAMISEN TULOKSET

Tässä osiossa on kuvattu Loton myynnin mallintamisen tulokset. Kuvattuna on muodostettujen regressiomallien identifiointiprosessi sekä luodut myynnin regressiomallit ja niiden tulkinta. Lopuksi on esitetty regressiomallien diagnostinen testaus ja sen tulokset.

Loton myyntiä mallinnettiin kahden toisistaan poikkeavan ideologian avulla. Tarkoituksena oli identifioida kaksi matemaattiselta lähtökohdaltaan toisistaan poikkeavaa mallia ja tarkastella, eroaako näillä saadut tulokset merkittävästi toisistaan, niin regressioanalyysin kuin simulointien osalta. Loton kestojen myynneille muodostettiin tavanomaiset usean selittäjän regressiomallit kaavassa (10) esitetyn periaatteen

mukaisesti. Tämän lisäksi kestoille muodostettiin logaritmoitu usean selittäjän regressiomalli. Tämä toteutettiin, sillä odotusarvona oli että esimerkiksi Loton potti vaikuttaa Loton myyntiin hyvin epälineaarisesti. Logaritmoidun usean selittäjän regressiomallin perusta on esitetty kaavassa (16).

Loton myynnille laadittiin kaksi poikkeavan ideologian mallia, jotta itse regressioanalyysin ja etenkin simulointien tuloksiin voitaisiin paremmin luottaa. Mikäli molemmilla regressiomallien rakenteilla saavutettaisiin samansuuntaisia tuloksia, voitaisiin niiden luotettavuutta pitää parempana.

7.1 LOTON MYYNNIN MALLIEN IDENTIFIOINTI

Loton myynnin mallintamisprosessi aloitettiin tarkastelemalla kansainvälisesti julkaistuja tutkimuksia ja havainnoimalla niissä lottopelien myynnin kannalta merkittäväksi nousseita selittäjiä. Tämän lisäksi kävin useita keskusteluja Veikkaus Oy:n eri asiantuntijoiden kanssa siitä, mitä selittäjiä malleihin kannattaisi pyrkiä sisällyttämään.

Loton myynnin mallintamisen suurimpana haasteena on Loton myyntiin vaikuttavien tekijöiden erittäin suuri määrä. Kaikkia myyntiin vaikuttavia tekijöitä ei voitu malleihin sisällyttää, mikä myös mallien identifiointivaiheessa tiedostettiin. Selittäjiä, joita malleihin ei tässä yhteydessä saatu sisällytettyä, olivat muun muassa markkinointitoimenpiteet ja säätila. Regressiomallien tulosten perusteella voitiin havaita, että Lotossa on havaittavissa inflatoitumisen piirteitä potin vaikuttavuuden osalta. Vastaavia kokemuksia on myös muualla maailmasta lottopelien osalta. Loton pottien inflatoitumisen vaikutusta ei onnistuttu luotettavasti regressiomalleihin sisällyttämään.

Regressiomallien identifiointiprosessin aikana testattiin noin kuudenkymmenen eri muuttujan vaikutusta pelaamiseen. Lopullisiin malleihin sisällytettiin 29 eri muuttujaa.

Koska työn pääasiallisena tavoitteena oli rakentaa simulaattori, jonka avulla Loton vuositason myynnin jakaumia voidaan tutkia, ei simulaattoria varten muodostetuissa regressiomalleissa liikaa varottu selittäjien välistä multikollineaarisuutta. Multikollineaarisuus nousi ongelmaksi lähinnä kertapelien myyntien mallinnosten yhteydessä, joten niille laadittiin erikseen vielä regressiomallit, joiden osalta multikollineaarisuus pyrittiin pitämään mahdollisimman alhaisena. Tämä toteutettiin, jotta Loton myyntiin vaikuttavia tekijöitä pystyttiin mahdollisimman luotettavasti arvioimaan myös kertapelien osalta.

7.2 LOTON MYYNNIN MALLINTAMISEN TULOKSET

Taulukossa 8 on esitetty laaditut regressiomallit parametreineen ja tilastollisine tunnuslukuineen. Taulukon ylintä otsikkoriviä luetaan seuraavasti: ensimmäinen numero

kertoo selitettävänä olevan myynnin keston ja toinen numero on järjestysnumero selitettävälle mallille. Järjestysnumero 1 kuvaa tavanomaista regressiomallia, kun taas järjestysnumero 2 kuvaa logistisen regression avulla laadittua myynnin mallia. Esimerkiksi malli 3-1 kuvaa kolmen viikon myynnin mallia tavanomaisella regressiolla toteutettuna.

Ensimmäinen regressiomalli on tavanomainen usean selittäjän regressiomalli ja sen periaate on kuvattu edellä kaavassa (10). Logaritminen regressiomalli on muotoa:

$$\ln(y_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i \rightarrow \quad (16)$$

$$y_i = e^{\beta_0} e^{\beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}} e^{\varepsilon_i} \quad (17)$$

Toinen malli identifioitiin, koska myynnin riippuvuus esimerkiksi potista ei selkeästi ole lineaarinen, vaan omaa epälineaarisia piirteitä. Toisen mallin haluttiin myös vahvistavan luotujen regressiomallien tulkintaa, sekä antavan lisäuskottavuutta toteutettaville simulaatioille.

Regressiomalleja 1 & 2 laadittaessa ei ole juurikaan kiinnitetty huomiota selittäjien multikollineaarisuuteen. Tätä perustellaan sillä, että näillä malleilla on pääsääntöisesti tarkoitus mallintaa kestojen myyntiä mahdollisimman tarkasti simulointia varten. Regressiokertoimien tulkinta ei ollut ensisijaisena tarkoituksena, toki siihenkin pyrittiin. Selittäjien välisen multikollineaarisuuden havaittiin olevan ongelmallista kertaerien malleilla 1-1 ja 1-2. Näin ollen kertaerille muodostettiin vielä toiset regressiomallit 1-1M ja 1-2M, joiden avulla pystyttiin luotettavammin tarkastelemaan kertaerien myyntiin vaikuttavia tekijöitä. Malli 1-1M oli kaavan (10) ja malli 1-2M kaavan (16) mukainen.

Muodostettujen regressiomallien tulokset on esitetty taulukossa 8. Taulukkoa muodostettaessa yksikkönä on ollut yksi miljoona. Taulukon jälkeen on kuvattu muodostetuissa regressiomalleissa käytetyt selittäjät sekä niiden myyntivaikutuksen tulkinta ja suuruusluokka.

Selittäjä/malli	ind.	1-1	2-1	3-1	5-1	10-1	1-2	2-2	3-2	5-2	10-2	1-1M	1-2M
Vakiotermi	0	1,202	0,069	0,044	0,165	-0,304	14,881	11,687	11,339	12,547	12,357	0,977	14,898
Pottityyppi	1		0,005				0,078	0,019					
Pottityyppi^2	2	0,134											
Rivihinta €	3	3,945	0,159	0,088			0,614	0,642	0,514			2,586	0,615
Potti_M€(PTS)	4	1,847	0,004	0,008	0,022			0,023	0,039	0,036			0,096
PTS^2	5	-0,588					0,013					0,135	
PTS^3	6	0,077											
PTS^5	7		3,E-06										
Potin suhde edelliseen	8						0,028						
Hinnankorotus seur. kierros (0/1)	9			0,045	0,344	4,456			0,238	0,496	1,842		
Arvotaan ylim. Potti (0/1)	10	0,402	0,007	0,004	0,018		0,074	0,034		0,030		0,441	0,078
Myynti_1vko_lag1	11	0,092					0,022					0,108	0,014
Myynti_1vko_lag2	12						0,007						0,006
Myynti_1vko_lag5	13				-0,008					-0,014	-0,039		
Trend (vko datan alusta)	14	-0,002	4,E-05	3,E-05			-2,E-04	2,E-04	2,E-04				-2,E-04
Arvonta sunnuntaina(0/1)	15					0,339					0,240		
Kierroksen ajankohta (5/13)	16	-0,284					-0,035					-0,257	-0,043
Kierroksen ajankohta (6/13)	17	-0,358					-0,057					-0,346	-0,056
Kierroksen ajankohta (7/13)	18	-0,442					-0,064					-0,387	-0,067
Kierroksen ajankohta (8/13)	19	-0,306					-0,045					-0,297	-0,050
Myynti_2vko_lag2	20		0,203					0,830					
Myynti_3vko_lag3	21			0,284					1,505				
Myynti_5vko_lag5	22				0,705	1,769				1,233	2,452		
Myynti_10vko_lag10	23				0,020	0,277				0,030	0,195		
Myyntipaikat tuhatta	24											0,554	
Pottityyppi-1(0/1)	25											0,467	0,057
Pottityyppi-2(0/1)	26											0,572	0,086
Pottityyppi-3(0/1)	27											1,118	0,162
Pottityyppi-4(0/1)	28											2,554	0,310
Pottityyppi-5(0/1)	29											2,730	0,361
RMSE		0,334	0,012	0,011	0,037	0,415	0,048	0,047	0,058	0,064	0,308	0,399	0,049
R^2		0,943	0,838	0,801	0,764	0,469	0,936	0,834	0,799	0,750	0,421	0,919	0,933
adj(R^2)		0,941	0,835	0,798	0,759	0,462	0,934	0,831	0,796	0,746	0,412	0,916	0,930
F		546	295	269	216	89	484	336	320	201	58	320	361

Taulukko 8: Muodostetut regressiomallit ja niiden tilastolliset tunnusluvut

Vakiotermi (ind. 0)

Tulkinta vakiotermiä osalta on haasteellinen. Esimerkiksi mallin 1-1M mukaan vakiotermiä osuus myynnistä olisi noin 1M€. Tämän lisäksi tulee vielä muiden selittäjien minimitasoa vaikutus, minkä senkin voidaan katsoa tietyllä tavalla vakioista. Vakiotermiä nojalla ei siis suoraan voida sanoa, paljonko myynnistä on riippumatonta muista selittäjistä, ns. tapapelaamista. Vakiotermiä voidaankin katsoa olevan matemaattinen vaatimus ja sille ei sen laajempaa tulkintaa tässä tehdä.

Pottityyppi ja pottityyppi^2 (ind. 1-2)

Muuttuja pottityyppi kuvaa sitä, kuinka moneen edelliseen kierrokseen ei ole Loton päävoittajaa löytynyt (0, 1, 2, ...). Muuttujalla voidaan havaita olevan selkeä vaikutus myyntiin yhden ja kahden viikon kestoissa. Tulosta voidaan pitää odotettuna, sillä mitä useampana viikkona Loton potti jää jakamatta, sitä enemmän ihmisten voidaan olettaa Lotosta kiinnostuvan. Muuttujan voidaan katsoa tukevan pottimuuttujaa, vaikkakin ne korreloivat suhteellisen voimakkaasti keskenään.

Kertapeliin osalta pottityyppimuuttujan vaikutusta pelaamiseen on käsitelty tarkemmin seuraavassa osiossa: pottityyppien I/O muuttujat. Kahden viikon kestoilla viikkotason myyntivaikutus vaikuttaisi olevan suuruusluokkaa 5-25k€.

Pottityyppien I/O muuttujat (ind. 25-29)

Tämä selittäjä indikoi I/O muuttujan avulla sitä, kuinka moneen kierrokseen ei Lotossa ole löytynyt päävoittoa, aivan kuten edellä kuvattu pottityyppimuuttuja. Selittäjää on hyödynnetty alhaisen multikollineaarisuuden malleissa 1-1M ja 1-2M, jotta pottityypin vaikutus kertapeliin pelaamiseen saataisiin mahdollisimman hyvin esille. Jälkikäteen ajateltuna pottityyppi olisi ehkä ollut järkevää ilmaista I/O muuttujien avulla myös muiden regressiomallien yhteydessä.

Pottityypin voidaan havaita kasvattavan pelaamisintoa siinä missä itse Loton absoluuttisen potinkin. Kierrostaolla voidaan havaita, että pottityypin vaikutus kertapelaamiseen on noin 400k€ - 3M€. Merkitys on siis erittäin suurta varsinkin tilanteessa, jossa Loton potti on jäänyt useamman viikon jakamatta. Tulosta voidaan pitää luonnollisena ja odotettuna, sillä mitä useammin potti jää jakamatta, sitä enemmän se herättää kiinnostusta niin mediassa kuin pelaajissakin.

Vertailtaessa pottityyppimuuttujan vaikutusta kertapelaamiseen ja kestopelaamiseen voidaan havaita, että pottityyppi vaikuttaa selvästi voimakkaammin kertapelaamiseen. Tulosta voidaan pitää suhteellisen odotettuna, sillä kertapelaamisen voidaan katsoa olevan enemmän niin sanottua herätepelaamista, kun taas kestopelaaminen on enemmänkin tapapelaamista. Mikäli Lotossa ei useampaan kierrokseen päävoittajia löydy, voisi tämän kuvitella kasvattavan nimenomaan Loton herätepelaamista.

Rivihinta (ind. 3)

Regressiomallien mukaan rivihinnalla on positiivinen vaikutus niin yhden, kahden kuin kolmen viikon kestopelien myyntiin. Tulkintana voidaan suoraviivaisesti siis pitää, että hinnankorotukset ovat kasvattaneet euromääräistä pelaamista etenkin lyhyemmillä kestoilla. Ainakin osittaisena syynä tähän voidaan olettaa olevan se, että osa pelaajista on säilyttänyt pelaamansa rivimääränsä ennallaan, vaikka pelin hinta on noussut. Hinnankorotuksen kierrostaon myyntivaikutukseksi voidaan arvioida kertapeliin osalta noin 250k€/0,1€ hinnankorotus. Kahden ja kolmen viikon kestopeleillä vastaava luku on noin 7-20k€/0,1€ hinnankorotus.

Potti, Potti², Potti³M€, Potti⁵ (ind. 4-7)

Loton pottia voidaan pitää kaikista luonnollisimpana ja samalla merkittävimpana selittäjänä Loton pelaamisen määrälle. Suurempien lottopottien aikana voi havaita että Loton pelaaminen kasvaa etenkin kertapeliin osalta hyvinkin voimakkaasti ja epälineaarisesti. Tarkasteltaessa taulukkoa 8 huomataan Loton potin olevan erittäin

merkittävä selittäjä kaikilla muilla paitsi 10 viikon kestoilla. Tulosta voidaan pitää tietyllä tapaa odotettuna, sillä 10 viikon pelaajien voidaan katsoa olevan niitä Lotolle uskollisimpia asiakkaita, joiden pelaamisen Loton potin suuruudella ei juuri ole merkitystä. Vastaavankaltaisia tuloksia potin merkityksestä pelaajille on saatu myös askaskyselyiden nojalla.

Potilla huomataan olevan selvää epälineaarista vaikutusta pelaamiseen, mitä voidaan pitää odotettuna. Esimerkiksi logaritmisessa regressiomallissa kertapelien pottiselittäjä on toisessa potenssissa, mikä voidaan tulkita niin että potilla todella on voimakkaasti epälineaarinen vaikutus pelaamiseen.

Potin vaikutus pelaamiseen kertapeleillä kierroksella on noin 1-10M€ ja vastaavasti kestopeleillä suuruusluokkaa 10-200k€. Tästä huomataan selkeästi se, että Loton potti vaikuttaa erityisesti kertapelaajiin ja maltillisemmin kestopelaajiin. Vaikuttaisi siltä, että mitä pidemmästä kestoista on kysymys, sitä vähäisempi suhteellinen vaikutus potin suuruudella on kyseisen keston pelaamiseen. Absoluuttisina euroina vaikuttaisi kuitenkin esimerkiksi siltä, että viiden viikon kesto reagoi euromääräisesti voimakkaammin Loton pottiin kuin kolmen viikon kestopelaaminen.

Potin suhde edelliseen (ind. 8)

Potin suhde edelliseen kuvaa tarkasteltavan kierroksen lottopotin suhdetta edellisen kierroksen lottopottiin. Muuttujalla oli tarkoitus tutkia sitä, vaikuttaako esimerkiksi potin nopea suhteellinen kasvu Loton pelaamiseen. Tietyissä tilanteissa voisi olettaa, että Loton potin nopea kasvaminen vaikuttaisi positiivisesti esimerkiksi kertapelaamiseen.

Logaritmisesta kertapelimallin 1-2 mukaan voidaan tulkita, että potin suhteella edelliseen olisi lievää positiivista vaikutusta pelaamiseen. Tätä selittäjää voidaan kuitenkin tulosten valossa pitää hieman kyseenalaisena, sillä muissa regressiomalleissa selittäjä ei noussut tilastollisesti merkittäväksi. Selittäjä kuitenkin pidettiin logaritmisessa kertapelimallissa mukana, koska sillä on tiettyjä edellä kuvattuja loogisia perusteita tilastollisten perusteiden lisäksi.

Seuraavalla kierroksella hinnankorotus (ind. 9)

Tulevalla kierroksella tapahtuva hinnankorotus vaikuttaa pelaamiseen merkittävästi eritoten pidemmillä kestoilla. Syynä tähän on se, että pelaajilla on ennen hinnankorotusta ollut mahdollista pelata vanhalla hinnalla uuteen peliin. Tämä on johtanut siihen, että hinnankorotusta edeltäneellä viikolla pitkät kestot, eritoten 10 viikon kesto, ovat kasvattaneet suosiotaan merkittävästi.

Kertapelien ja kahden viikon kestopelien pelaamiseen ei hinnankorotuksella voida katsoa olevan selkeää vaikutusta. Kolmen viikon kestopeleihin vaikutuksen suuruus-

luokka on kierrostaalla noin 50k€, 5 viikon kestoihin noin 350k€ ja 10 viikon kestoihin jopa neljä miljoonaa euroa.

Arvotaan ylimääräinen potti eli lisäarvonta (ind. 10)

Kyseessä on I/O muuttuja, jolla indikoidaan sitä, että varsinaisen lottoarvonnan lisäksi kierroksella on ilman pelaajilta perittävää lisämaksua arvottu ylimääräinen 7 oikein - tulos. Pelaaja on voittanut lisäarvonnassa, mikäli hänellä on ollut kyseisessä arvonnassa 7 oikein -tulos. Lisäarvontaan ovat olleet oikeutettuja kaikki Loton varsinaiseen arvontaan osallistuneet pelirivit.

Lisäarvonnalla voidaan nähdä olevan selkeästi positiivinen vaikutus Loton myyntiin. Lisäarvonta vaikuttaa eniten kertapeliin pelaamiseen. Regressiomallien mukaan vaikutus myyntiin kertapeleillä on noin 400k€:n suuruusluokkaa. Kestopelien osalta lisäarvonnat näyttäisivät vaikuttavan eritoten 2 ja 5 viikon kestopelien pelaamiseen. 3 viikon kestopelaamiseen vaikutuksen voidaan katsoa olevan hieman tulkinnanvaraisista. Kestopelaamisen vaikutuksen suuruusluokan voidaan katsoa olevan 5-20k€:n suuruusluokkaa, eli huomattavasti vähemmän kuin kertapeliin tapauksessa.

Kertapeliin ja kestopeliin vaikutusten suuruusluokkien suhdetta voidaan pitää tietyssä mielessä odotettavissa olevana. Lisäarvonnat on järjestetty useimmiten pelaajien näkökulmasta satunnaisilla viikoilla. Näin ollen on helppo kuvitella, että lisäarvonnan merkittävänä lisäarvona kokevat pelaajat oletetusti osallistuvat kyseiselle kierrokselle pelaamalla Lottoa nimenomaan kertapeliin eikä kestonä.

Yhden viikon kestojen myynnin viivästetyt selittäjät (Myynti_1vko_Lag 1,2,5) (ind. 11-13)

Selittäjä kuvaa kertapeliin myyntiä edellisten kierrosten osalta. Selittäjän avulla pyritään tutkimaan sitä, kuinka paljon tiettyjen kierrosten positiivinen myynti kertapeliin osalta heijastuu tulevien kierrosten pelaamiseen. Oletusarvona on, että mikäli pelaaja aktivoituu Lottoa pelaamaan esimerkiksi suuren potin myötä, on todennäköisempää, että pelaaja pelaa Lottoa myös tulevilla kierroksilla.

Regressiotulosten nojalla voidaan sanoa, että kertapeliin myynti heijastuu positiivisesti myös tuleville kierroksille. Tulosta voidaan pitää luonnollisena, sillä satunnaisen pelaajan (todennäköisesti kertapelaaja) aktivoiduttua pelaamaan Lottoa voisi ajatella, että hän helpommin jatkaa pelaamista myös tulevilla kierroksilla. Tuloksista voidaan havaita, että edellisen kierroksen kertapeliin pelaaminen kasvattaa pelaamista selkeästi enemmän kuin kahden viikon takaisten kertapeliin pelaaminen. Tavallisen regressiomallin tuloksista oli havaittavissa myös mielenkiintoinen, tosin hieman kyseenalainen tulos, jonka mukaan viiden viikon takainen kertapeliin pelaaminen vähentäisi kuluvan kierroksen viiden ja kymmenen viikon kestojen pelaamista. Kymmenen viikon osalta tulokset oli nähtävissä ainoastaan logaritmisessa mallissa. Tulosta voidaan pitää yllättävänä, mutta sille on olemassa varauksin myös loogisia perusteita. Jos

esimerkiksi viisi viikkoa sitten olisi Lotossa ollut suuri potti, olisi tämä johtanut kasvaneeseen kertapeliin ja viiden viikon kestojen pelaamiseen. Kun viisi viikkoa tästä on kulunut, päättyvät viisi viikkoa sitten pelatut kestot ja jos Lotossa ei ole merkittävää pottia tarjolla, niin osa kestopelaajista jää oletettavasti uusimatta. Tuloksella on siis olemassa loogisia perusteita, tosin mitenkään yksiselitteisinä niitä ei voida pitää.

Trendi (ind. 14)

Trendi-muuttuja kuvaa sitä, onko ajan kuluessa havaittavissa säännönmukaista muutosta Loton pelaamisessa. Käytetty trendi-muuttuja on muotoa 1, 2, 3... 408 ja kattaa koko regressiomallien muodostuksessa käytetyn datan.

Tulosten nojalla trendimuuttujan perusteella voisi sanoa, että kertapeliin pelaaminen on lievässä laskussa, kun taas lyhyiden kestojen (2 ja 3 viikkoa) kestopelaaminen hienohkossa nousussa. Kertapeliin osalta trendi-muuttujan voisi ajatella kuvaavan varauksin Loton pottien inflatoitumista. Toisaalta voisi myös kuvitella, että ajan saatossa Veikkaus on lanseerannut uusia rahapelejä, joiden vaikutus saattaisi näkyä kertapeliin osalta laskevana trendinä, kun pelaajat ovat siirtäneet osan pelaamisestaan uusiin peleihin. Lyhyiden kestopeliin osalta on vaikea sanoa, miksi trendi-muuttuja saa niiden osalta nousevan arvon. Voisi ehkä kuvitella, että vaikka lyhyemmät kestot ovat olleet pelaajien pelattavissa koko tarkasteltavan ajankohdan yli, ovat ne kuitenkin vielä suhteellisen pienessä roolissa Loton pelaajien keskuudessa. Näin ollen lyhyillä kestoilla voisi kuvitella Loton asiakkaiden joukossa olevan pelaajia, jotka eivät ole vielä lyhyitä kestoja niin sanotusti löytäneet.

Kertapeliin osalta vaikuttaisi siltä, että vuositasolla niiden pelaaminen trendinomaisesti vähenisi noin 80k€ yhtä kierrosta kohden. Vastaavasti lyhyiden kestojen osalta vuositason kierroskohtainen kasvu on hyvin vähäistä, joitakin tuhansia euroja.

Sunnuntaiarvonta (ind. 15)

Sunnuntaiarvonta muuttuja on I/O muuttuja, jolla kuvataan niitä kierroksia jolloin Loton arvonta on järjestetty sunnuntaisin. Muuttujan vaikutusta myyntiin on hieman haasteellinen tulkita, sillä regressiotulosten mukaan vaikutusta olisi ainoastaan kymmenen viikon kestopeliin pelaamiseen. Regressiomallien mukaan sunnuntaiarvonnan vaikutuksen suuruusluokka kymmenen viikon kestopelaamiseen olisi suuruusluokaltaan noin 350k€. Mahdollisia selittäjiä sunnuntaiarvonnan positiiviseen vaikutukseen ovat muun muassa se, että sunnuntaiarvonnat on usein järjestetty yksittäisinä sunnuntaina. Tällaisessa tapauksessa Loton kierroksella on ollut kahdeksan pelipäivää, normaalin seitsemän sijaan. Toisaalta Loton arvonnat järjestettiin sunnuntaisin säännönmukaisesti kesällä 2007, jolloin edellä kuvattu logiikka ei ole pätevä. Sunnuntaiarvonta saattaa poikkeavuudellaan myös herättää ihmisten 10 viikon kestopelaamista ja näin ollen saada ihmiset varmistamaan osallistumisensa Loton tuleviin arvontoihin pitkän kestopelin avulla.

Kierroksen ajankohta (ind. 16-19)

Kierroksen ajankohta on I/O muuttuja jonka avulla vuosi jaettiin neljän kierroksen mittaisiin periodeihin, eli pääsääntöisesti vuotta kohden tuli tasan 13 neljän kierroksen periodia. Kierroksen ajankohta muuttujan avulla voidaan huomata että vuodenajalla vaikuttaisi olevan kesän osalta negatiivinen vaikutus Loton kertapeliin myyntiin. Tulosta voidaan pitää luonnollisena, sillä kesäaikana ihmiset lomailevat ja heidän viikkorytminsä poikkeaa usein selkeästi tavanomaisesta. Kesäaikana Veikkauksen urheilupelien tarjonta on usein muuta vuotta vähäisempää, minkä voidaan olettaa vaikuttavan siten että tietty osa ihmisistä ei kesäaikana pelaa niin säännöllisesti Veikkauksen pelejä. Kestopelaamiseen vuodenajalla ei kuitenkaan näyttäisi olevan merkittävää vaikutusta. Tulos on sinällään odotettavissa oleva, sillä kestopelaaminen on enemmänkin tapapelaamista ja näin ollen voisi kuvitella, ettei vuodenajalla ole pelaajalle niin merkittävää vaikutusta.

Vuodenajalla vaikuttaisi olevan kertapelaamiseen pelaamista vähentävä kierrosta-
sol- la noin 250 - 400k€:n suuruusluokaltaan oleva vaikutus. Vaikutus osuu touko-
elokuulle ja on voimakkaimmillaan keskikesällä. Kestopeleihin ei vuodenajalla näyt-
täisi olevan merkittävää vaikutusta, kuten edellä jo todettiin.

Viivästetyt kestopyyntien selittäjät (Myynti_2vko_lag2 jne.) (ind. 20-23)

Viivästetyillä kestopyyntien selittäjillä tarkoitetaan tietyn keston osalta edeltänei-
den kierrosten kestopyyntien vaikutusta kuluvaan kierroksen kestopelaamiseen. Esi-
merkiksi selittäjä myynti_2vko_lag2 tarkoittaa kahden kierroksen takaisen kahden
viikon kestopyyntien vaikutusta kuluvaan kierroksen kahden viikon kestopelien pe-
laamiseen. Mikäli esimerkiksi viisi viikkoa sitten viiden viikon kestoja pelattiin miljoon-
alla eurolla tarkoittaa tämä sitä, että kuluvalle kierroksella päättyy miljoonan euron
edestä viiden viikon kestopelien. Pelaajat usein uusivat kestopelinsä, joten päättyvi-
en kestojen euromäärä indikoi vahvasti kuluvaan kierroksen kestopelaamisen suuruus-
luokkaa.

Edeltävien kierrosten kestopelien myynnin voidaan havaita vaikuttavan merkittävästi
kaikkien kestojen pelaamiseen. Kahden viikon kestojen osalta merkittäväksi selittä-
jäksi nousee kahden viikon takainen kahden viikon kestopelien myynti. Vastaava ha-
vainto voidaan tehdä myös kolmen viikon kestopelien osalta. Hieman yllättävästi tu-
loksista voidaan havaita, että viiden viikon kestopelaamiseen vaikuttaa selvästi myös
kymmenen viikon kestopelaaminen ja vastaavasti kymmenen viikon kestopelaami-
seen viiden viikon kestopelaaminen. Tulosta voidaan tulkita siten, että mahdollisesti
pelaajat vaihtelevat pitkien kestojen välillä, eli välillä he valitsevat viiden ja välillä
kymmenen viikon kestopelin.

Vaikutusten suuruusluokkien perusteella voidaan havaita, että mitä pidemmästä kes-
tosta on kysymys, sitä vahvemmin edeltävien syklien vastaavan kestopelin pelaami-

nen kyseisen keston myyntiin vaikuttaa. Tulosta voidaan pitää luonnollisena, sillä pidempien kestojen voidaan olettaa olevan vahvemmin tapapelaamista, jolloin pelaajat uskollisesti uusivat kestojaan niiden voimassaolon päätyttyä. Kymmenen viikon kestopelaamisen osalta voidaan havaita, että hieman yllättäen viiden viikon takaisen viiden viikon kestopelin myynnin vaikutuksen kerroin on suurempi kuin kymmenen viikon keston vastaavan kertoimen. Tulos ei kuitenkaan suoranaisesti tarkoita sitä, että euromääräisesti viiden viikon kestolla olisi suurempi vaikutus kymmenen viikon kestopelaamiseen, sillä euromääräisesti 10 viikon kestopelaaminen on selvästi suositumpaa kuin viiden viikon kestopelaaminen (ks. taulukko 6).

Myyntipaikkojen lukumäärä (ind. 24)

Myyntipaikkojen lukumäärä kuvaa kullakin kierroksella aktiivisten Veikkauksen myyntipaikkojen lukumäärää. Myyntipaikkojen lukumäärällä voidaan varauksin katsoa olevan lievä positiivinen vaikutus Loton kertapelien pelaamiseen. Myyntipaikkojen lukumäärä nousi merkittäväksi selittäjäksi ainoastaan mallissa 1-1M, joten tätä tulosta ei voida täysin yksiselitteisenä pitää.

Tulosta voidaan mielestäni tulkita niin, että Lotto on peli, jonka pelaamisen eteen ihmiset ovat valmiita näkemään hieman jopa vaivaa. Näin ollen myyntipisteiden kasvattaminen ei välttämättä merkittävästi lisää Loton pelaamista. On kuitenkin mahdollista, että kertapelaaminen kasvaa jonkin verran myyntipaikkojen lukumäärän myötä. Tätä tulosta voidaan pitää luonnollisena, sillä oletettavaa on, että mitä useampi myyntipaikka Lotolla on, sitä enemmän se tukee niin sanottua satunnaista kertapelaamista.

Seuraavassa on käsitelty muodostettujen regressiomallien tilastollisia ominaisuuksia. Tämän lisäksi on tarkasteltu mallien rakenteita niissä käytettyjen selittäjien osalta.

Identifioitujen regressiomallien tilastolliset ominaisuudet

Taulukossa 8 on esitetty muodostettujen regressiomallien tilastollisia tunnuslukuja. Tilastollisista tunnusluvuista voidaan havaita, että selitysasteet ovat sitä korkeammat, mitä lyhyemmästä kestoista on kysymys. Erityisen hyvin regressiomallit kykenevät selittämään kertapelien myynnin vaihteluita, mistä kertoo että niin mallien selitysaste R^2 ja korjattu selitysaste $\text{adj}(R^2)$ saavat yli 0,93 arvot (poislukien alhaisen multikollineaarisuuden malli 1-1M). Myös 2, 3 ja 5 viikon kestojen selitysasteet ovat korkealla tasolla ja nekin saavat yli 0,74 arvot niin tavallisen kuin logistisen regressiion malleilla. 10 viikon kestojen selitysaste jää selkeästi edellä mainittuja alhaisemmaksi, mutta saa sekin yli 0,41 arvot molemmilla regressiomalleilla.

Odotettavissa oli, että kertapelaamisen selittämisessä onnistutaan parhaiten. Perusteluina tälle voidaan katsoa olevan muun muassa se, että kertapelaaminen riippuu kaikista voimakkaimmin tarjolla olevasta potista sekä siitä, kuinka monta kierrosta Loton päävoitto on jäänyt jakamatta. Mallien selityskyvystä voidaan tehdä johtopääh-

tös, että mitä lyhyemmästä kestopaasta on kysymys, sitä enemmän sen pelaaminen riippuu niin sanotuista ulkoisista tekijöistä, kuten Loton potista. Pidemmät kestot puolestaan ovat enemmänkin tapapelaamista. 10 viikon keston myynnin mallintaminen oli hyvin haastavaa, mutta senkin osalta onnistuttiin myyntityninvaihteluiden estimoinnissa kohtuullisesti.

Vertailtaessa tavallisen ja logistisen regressiomallin selityksasteita huomataan niiden olevan kullekin kestolle hyvin samankaltaisia (poislukien mallit 1-1M ja 1-2M). Näin ollen pelkästään selityksasteiden tai tilastollisten tunnuslukujen avulla ylipäättään ei voida tehdä johtopäätöksiä siitä, kumpi regressiomalli olisi Loton myyntityninv mallintamisessa toimivampi.

Vertailtaessa matalan multikollineaarisuuden malleja 1-1M ja 1-2M voidaan tehdä mielenkiintoinen havainto siitä, että logistinen regressiomalli omaa suhteellisen selvästi korkeamman selityksasteen kuin tavallisen regression malli. Tulosta voidaan siinä mielessä pitää luonnollisena, että logistinen malli kykenee odotetusti paremmin mallintamaan Loton potin epälineaarista vaikutusta myyntiin, kun käytössä on ainoastaan yksi Loton potin suuruuteen sidottu muuttuja. Malleissa 1-1M ja 1-2M pyrittiin selittäjien multikollineaarisuus pitämään mahdollisimman alhaisena, mikä vaikutti osaltaan mallien selittäjien valintaan.

Identifioitujen regressiomallien rakenne

Yleisesti ottaen voidaan sanoa, että niin tavallisen kuin logistisen regression avulla muodostetut mallit omaavat rakenteellisesti hyvin paljon samoja selittäjiä. Tulosta voidaan siinä mielessä pitää odotettuna, että olisi ollut hyvin yllättävää, mikäli eri menetelmillä olisi saatu selkeästi ristiriitaisia tuloksia siitä mitkä tekijät Loton myyntiin eniten vaikuttavat.

Taulukosta 8 voidaan havaita että 5 viikon kestopaasten mallit ovat selittäjien osalta identtiset tavallisella ja logistisella regressiolla. Muiden kestopaasten osalta voidaan havaita, että erot mallien selittäjien välillä ovat suhteellisen vähäisiä. Tämä voidaan mielestäni tulkita siten, että mallien rakenteissa tuskin on ainakaan merkittävää säännönmukaista virhettä ja muodostettujen mallien avulla voidaan tehdä tulkintaa siitä, mitkä tekijät Loton pelaamiseen vaikuttavat.

7.3 REGRESSIOMALLIEN GRAFIIKKAA JA DIAGNOSTINEN TESTAUS

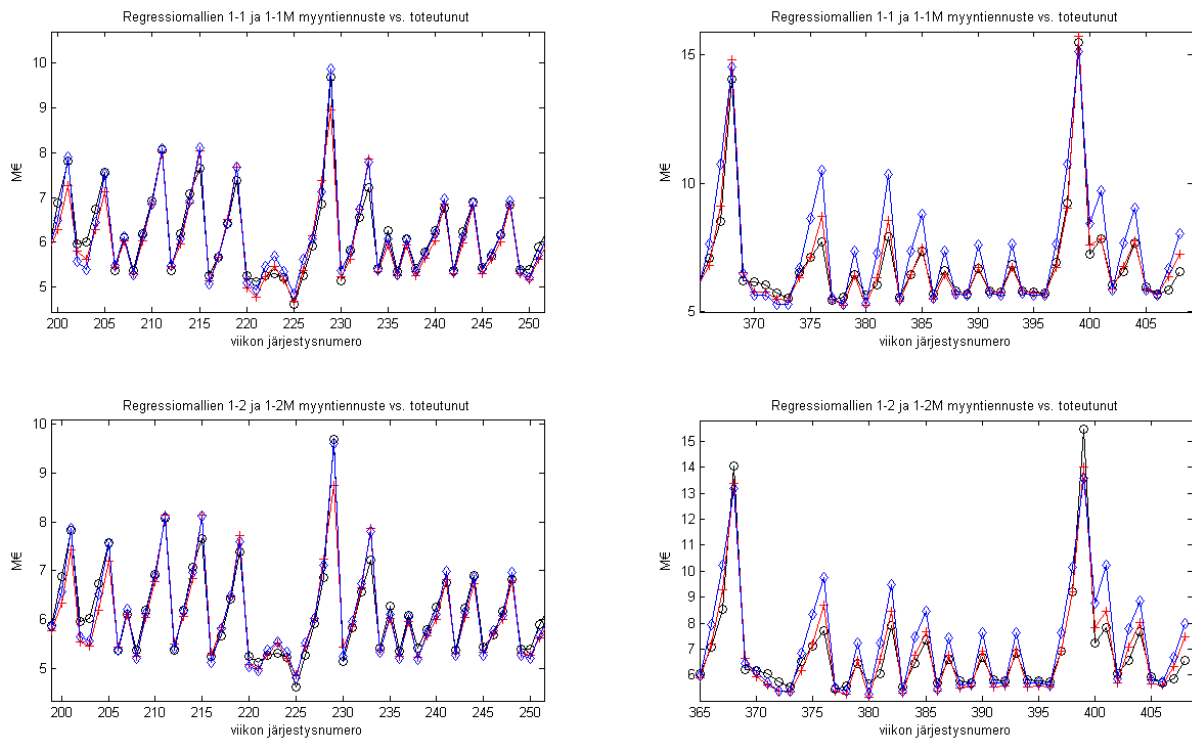
Tässä osiossa on tarkasteltu regressiomallien suoriutumista niin graafisten kuvaajien, kuin tilastollisten testienkin avulla. Kuvaajien ja tilastollisten testien avulla on tarkasteltu muun muassa sitä, kuinka hyvin muodostetut mallit selittävät myyntityninv vaihtelua, ovatko mallien residuaalit normaaleja ja homoskedastisia sekä onko residuaaleissa havaittavissa autokorrelaatiota.

7.3.1 REGRESSIOMALLIEN KUVAAJIA

Kuvissa 5 ja 6 on esitetty muodostettujen regressiomallien estimaattorit todellisia myyntilukuja vastaan. Kuvassa 5 on esitetty kertapeliin ja kuvassa 6 kestopeliin regressiomallien suoriutuminen. Todelliset myynnit on kuvattu mustan ja estimaattorit punaisen viivan avulla. Kuvassa 5 on sinisellä viivalla kuvattu malleja 1-1M ja 1-2M. Viikon järjestysnumero tarkoittaa tässä yhteydessä havainnon järjestysnumeroa estimointidatan alusta. Aineistossa oli yhteensä 408 eri havaintoa.

Kuvasta 5 voidaan havaita että kertapeliin osalta regressiomallit 1-1 ja 1-2 toimivat erittäin hyvin. Tämä oli odotettavissa, sillä kertapeliin osalta regressiomallien selityksasteet olivat huomattavan korkeat. Muodostettujen mallien välisestä suoriutumisesta on hyvin hankala sanoa näiden kuvien avulla mitään. Vaikuttaisi siltä, että malli 1 kykenee herkemmin ennustamaan poikkeuksellisen suuria myyntejä. Tämän lisäksi voidaan havaita, että mallinnosvirheiden (residuaalien) merkki tuntuisi hyvin usein olevan sama molemmilla malleilla. Eli jos toisella malleista saadaan myynteiksi enemmän kuin toteutunut myynti, niin vastaavaan tulokseen päädytään myös toisella mallilla. Tätä voidaan pitää suhteellisen odotettuna tuloksena, sillä olihan mallien selittäjät lähes yhtenevät. Kertapeliin mallien ennustuskyvystä voidaan varauksin vielä todeta, että näyttäisi siltä että datan loppupäässä mallit herkemmin yliarvioivat kertapeliin myyntiä. Tätä voidaan pitää yhtenä indikaattorina Loton potin vaikutuksen inflatoitumisesta, mitä regressiomalleihin ei saatu sisällytettyä.

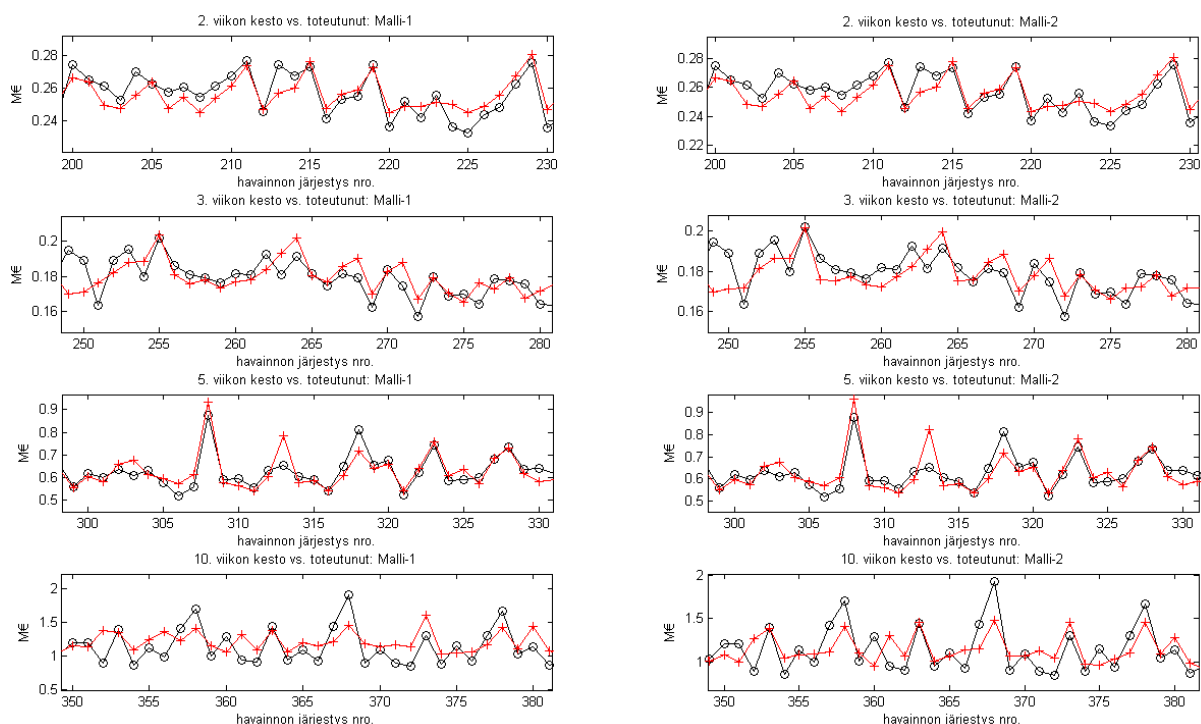
Mallit 1-1M ja 1-2M näyttäisivät toimivan paremmin datan keskivaiheilla (vasemman puoleiset kuvaajat). Silmämääräisesti ennustemallit toimivat erittäin hyvin. Oikeanpuoleisissa datan loppupään kuvissa mallit vaikuttaisivat yliarvioivan kierrosmyyntejä suhteellisen säännönmukaisesti. Tämä sama tulos voidaan havaita osiosta 7.3.2.3 löytyvistä kuvista 8 ja 9. Datat loppupään myyntien yliarviointi johtunee pääosin Loton pottien inflatoitumisesta ajan mittaan, mikä aiheuttaa estimaattoreiden ylisuuret ennusteet.



Kuva 5: Kertapeliin regressioiden mukaiset myynnit vs. toteutuneet myynnit

Kuvassa 6 on esitetty kestopelien regressiomallien suoriutuminen todellisia myyntejä vastaan. Kestopelien mallien selitysasteet olivat samaa suuruusluokkaa kestoilla 2,3,5, kun taas keston 10 selitysaste oli selkeästi alhaisempi. Tarkasteltaessa kuvaa 6 voidaan havaita että 2,3,5 viikon kestojen estimaattoreiden suoriutuminen vaikuttaisi olevan keskenään hyvin samankaltaista, kun taas 10 viikon keston mallit suoriutuu tehtävästään selkeästi heikommin.

Kestopelien osalta voidaan tehdä vastaava havainto kuin kertapeliin osalta. Vaikuttaa siltä, että mallien residuaalien merkki on hyvin usein sama. Tämän lisäksi silmäääräisesti voidaan havaita, että suuret ennustevirheet toteutuvat usein samanaikaisesti. Havaintoa voidaan kestopelienkin osalta pitää suhteellisen luonnollisina, sillä regressiomallien selittäjät ovat lähes yhtenevät myös kestojen malleissa.



Kuva 6: Kestopelien regressioiden mukaiset myynnit vs. toteutuneet myynnit

7.3.2 DIAGNOSTINEN TESTAUS

Tässä osiossa on kuvattu laadittujen regressiomallien diagnostinen testaus ja sen tulokset. Diagnostisten testien tueksi on usein liitetty kuvaaja, joka havainnollistaa regressiomallien ominaisuuksia.

7.3.2.1 Poikkeavat havainnot ja regressiokertoimien vakioisuus

Poikkeavat havainnot

Poikkeavien havaintojen olemassaoloa tutkittiin regressiomallien residuaalien avulla. Mallien residuaalit on kuvattu kuvissa 8 ja 9. Residuaalien aikasarjoista voidaan havaita, että tilastollisessa mielessä poikkeavia havaintoja mitä todennäköisemmin datasta löytyy.

Poikkeavien havaintojen tiedetään vaikuttavan regressioanalyysissä seuraavasti (Mellin; Lineaarinen regressioanalyysi): mallin valinta ja estimointi vaikeutuu, sekä tilastollinen päättely saattaa vääristyä. Poikkeavia havaintoja etsittäessä kiinnitettiin lähinnä huomiota siihen, vaikuttaisiko niiden poistaminen merkittävästi mallien selittäjien valintaan. Tämän lisäksi tiedostettiin myös se, että poikkeavat havainnot hankaloittavat omalta osaltaan tilastollista päättelyä Loton myyntiin vaikuttavista tekijöistä.

Poikkeavat havainnot etsittiin mallien residuaalidatasta Grubbsin testin avulla (Grubbs F. E; 1969), jonka jälkeen poikkeavat havainnot poistettiin ja regressiomallit

evaluoitiin uudelleen. Tätä proseduuria käytiin läpi, kunnes poikkeavia havaintoja ei enää löytynyt. Tämän prosessin aikana huomattiin, että poikkeavien havaintojen poistaminen ei merkittävästi vaikuta regressiomallien selittäjien valintaan. Näin ollen tässä mielessä mitään merkittävää uutta tietoa myynteihin vaikuttavista tekijöistä ei saatu. Tämän lisäksi huomattiin myös se, että todella usein poistettavaksi valikoituivat sellaiset havainnot, jolloin tarkasteltavan keston myynti oli ollut korkea. Tämä korostui etenkin kertapeliin osalta. Tämä aiheutti sen, että mallien selityskyky suuremmilla poteilla aleni merkittävästi, kun estimointidataa niiden osalta hävisi.

Koska poikkeavien havaintojen poistosta ei vaikuttanut olevan mitään todellista hyötyä mallien valinnan tai selityskyvyn parantamisen suhteen, päädyin ratkaisuun, jossa en poistanut poikkeavia havaintoja. Tätä perustellaan seuraavasti: Poikkeavien havaintojen poisto ei vaikuttanut juurikaan mallin valintaprosessiin, mutta heikensi mallien selityskykyä huomattavasti. Selityskyvyn heikkeneminen oli havaittavissa etenkin kierroksilla, jolloin tarjolla oli kohtalaisen suuri potti. Edellä mainittujen lisäksi voidaan myös varmasti todeta, että yhdessäkään selitettävän tai selittävän muuttujan arvossa ei ole satunnaisia virheitä, jotka saattaisivat poikkeavia havaintoja aiheuttaa. Loton myyntiin vaan vaikuttaa niin moni muukin asia kuin mitä näihin malleihin on kyetty sisällyttämään, että poikkeavien myyntien olemassaolo on vaan hyväksyttävä. Mitä enemmän niitä koittaa mallia evaluoidessa poistaa, sitä enemmän mallien kyky selittää Loton pelaamisen luonnetta kärsii.

Regressiokertoimien vakioisuus

Regressiokertoimien vakioisuutta testattiin Chow-testillä (Mellin; lineaarinen regressioanalyysi). Testi toteutettiin siten, että regressioparametrit määritettiin varsinaisen regressiomallien lisäksi myös ensimmäisen 250 havainnon perusteella (datassa 408 havaintoa). Näin saadun mallin ennustekykyä verrattiin Chow-testin mukaisesti koko datasta estimoituihin malleihin.

Taulukossa 9 on esitetty testin tulokset regressiokertoimien vakioisuudelle. Raja-arvo on määritetty viiden prosentin merkitsevyystason mukaan. Taulukosta voidaan havaita, että Chow-testin mukaan regressiokertoimet olisivat vakioita kestopelien malleilla. Kertapeliin osalta tulos on tiukempi ja ainoastaan malli 1-2 läpäisee tilastollisen testin, sekin tosin äärimmäisen niukasti.

Tulosta voidaan pitää odotettuna. Kertapeliin myynti on selkeästi eniten riippuvainen tarjolla olevasta potista. Käytännön havainnot ovat osoittaneet, että myös Loton potteihin kohdistuu inflaatiota, aivan kuten mihin tahansa muuhun tarjoukseen. Näin ollen on luontevaa, että ajan mittaan tietyn suuruinen potti menettää kiinnostustaan ihmisten mielikuvissa. Verrokkiajanjaksolla (havainto 309; kierros 6/2008) Lotossa tapahtui päävoittoluokan osuuden kasvattaminen 10% → 13,6% (36 prosentin kasvu päävoitoissa). Tämä kasvatti potteja 36 %, mutta oletettavasti kasvaneiden pottien vaikutus ei suoraan siirtynyt myynteihin potin regressiokertoimen mukaan. Näin ollen

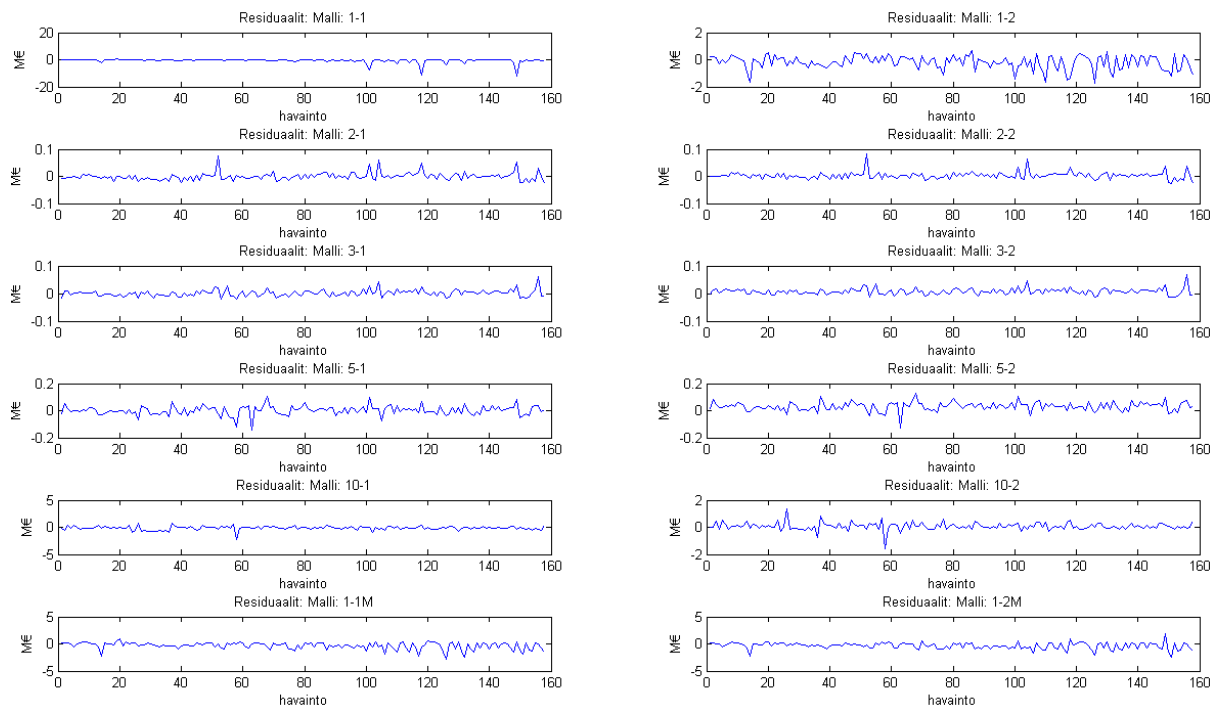
on odotettavaa, että ensimmäisen 250 havainnon perusteella estimoitu data hieman yliarvioi Loton myyntejä, etenkin kertapeliä osalta jälkimmäisessä verrokkidatassa. Tämä havainto voidaan vahvistaa myös kuvasta 7, jossa on kuvattu ensimmäisen 250 havainnon perusteella tehtyjen mallien ennustekyky jälkimmäiselle 158:lle havainnolle. Kuvaajista voidaan selkeästi nähdä, että kertapeliä osalta negatiiviset residuaalit ovat vallitsevia. Vastaavaa ilmiötä ei ole yhtä selkeästi havaittavissa muilla kestoilla. Residuaalien magnitudista voidaan havaita, että logaritmisella mallilla ei tule yhtä suuria ylilyöntejä ennusteisiin kuin tavanomaisella regressiomallilla.

Malli	1-1	2-1	3-1	5-1	10-1	1-2	2-2	3-2	5-2	10-2	1-1M	1-2M
F-testi	1,329	0,692	0,839	0,410	0,253	1,237	0,586	0,523	0,338	0,282	1,662	1,697
Raja-arvo	1,238	1,238	1,237	1,237	1,237	1,238	1,237	1,237	1,237	1,237	1,238	1,238

Taulukko 9: Regressiokertoimien vakioisuuden testaaminen

Kuvan 7 residuaaleista voi myös tehdä johtopäätöksen siitä, että mallien selityskyky on sitä parempi mitä lähempänä ollaan origoa, eli estimointidatan päätepistettä ja verrokkidatan alkupistettä. Lähes poikkeuksetta residuaalien vyö on kapeampi origon läheisyydessä. Tulosta voidaan pitää luonnollisena tässä tutkimustapauksessa. Mitä pidempi aika mallin parametrien estimointidatan päätepisteestä on kulunut, sitä enemmän maailma (ja Lotto pelin pelaaminen) on muuttunut ja mallin ennustekyky myynnin vaihteluiden suhteen heikkenee.

Parametrien vakioisuus oli yksi haasteellisimmista osista Loton myynnin mallintamisessa. On aivan selvää että 3M€ potti vuonna 2004 oli asiakkaille huomattavasti merkittävämpi kuin vuonna 2008. Osittain se johtuu siitä, että ajan myötä Loton potit ovat kasvaneet. Toisaalta myös inflaatio ja rahan arvon lasku on tehnyt tehtävänsä. Tätä inflatoitumisilmiötä, joka oletettavasti pääsääntöisesti selittyy päävoiton osuuden muutoksella, ei kyetty tässä työssä luotettavasti mallintamaan ja se jääkin tulevaisuuden tutkimuskohteeksi.



Kuva 7: Parametrien vakioisuus – regressiomallien residuaalit

7.3.2.2 Selittäjien multikollineaarisuus

Taulukossa 10 on kuvattu regressiomalleissa käytettyjen selittäjien keskinäiset korrelaatiot (Taulukossa 8 on listattu selittäjät indekseineen). Punainen väri indikoi positiivisesta ja sininen väri negatiivisesta korrelaatiosta.

Taulukosta voidaan havaita, että joidenkin selittäjien välillä vallitsee hyvinkin voimakas positiivinen korrelaatio. Odotettavasti kaikki pottiselittäjät sekä pottityyppiselittäjät omaavat vahvan positiivisen korrelaation keskenään. Yksittäisenä mielenkiintoisena havaintona voidaan huomata, että myyntipaikkojen lukumäärä (ind. 24) omaa negatiivisen korrelaation useiden selittäjien kanssa. Tämä selittyy ainakin osittain sillä, että havaintoaineistossa myyntipaikkojen lukumäärä on ollut lievässä laskussa, kun monien muiden selittäjien trendi on ollut ylöspäin. Selittäjien multikollineaarisuus on malleja laadittaessa tiedostettu ja se nousi merkittäväksi lähinnä malleilla 1-1 ja 1-2.

Taulukossa 11 on kuvattu laadittujen regressiomallien 1 & 2 matriisien kuntoisuusluvut multikollineaarisuuden osalta. Taulukko muodostettiin siten, että kunkin regressiomallin selittäjien korrelaatiomatriisien suurimman ja pienimmän ominaisarvon suhdetta tutkittiin. Menetelmä on kuvattu Mellinin luontomateriaaleissa: "Lineaarinen regressioanalyysi" (Mellin; lineaarinen regressioanalyysi). Taulukosta voidaan huomata, että multikollineaarisuus on selkeästi ongelmallisinta mallissa 1-1. Näin ollen tätä mallia ei kannata käyttää regressiokertoimien tulkinnessa, vaan kertapeliin myyntiin vaikuttavia tekijöitä arvioitaessa kannattaa tutkia ennemmin mallia 1-1M. Taulukosta 11 voidaan havaita, että regressiomalli 1-2M on yllättävästi jopa hieman multikol-

lineaarisempi kuin 1-2. Ero mallien välillä on kuitenkin hyvin pieni, eikä siihen tässä vaiheessa kannata kiinnittää sen suurempaa huomiota.

Taulukossa 11 on myös esitetty suurimman kuntoisuusluvun suhde muihin kuntouuslukuihin. Tämä kertaluokka kuvaa hyvin sitä, kuinka multikollineaarinen malli 1-1 oikeastaan on. Kuten edellä jo todettiin, multikollineaarisuus ei haittaa ennusteen arvoa laskettaessa, joten ennustemallia 1-1 voitiin hyvin käyttää vuosimyyntejä si-
muloitaessa.

IND.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
1	1,000	0,914	0,017	0,836	0,746	0,621	0,414	0,462	-0,020	-0,230	0,111	-0,044	-0,020	0,024	-0,006	0,083	-0,031	-0,031	0,075	0,030	0,017	0,004	-0,078	-0,058	0,116	0,462	0,519	0,408	0,284
2	0,914	1,000	-0,024	0,733	0,705	0,620	0,437	0,199	-0,025	-0,154	0,224	-0,007	-0,030	-0,018	-0,003	0,113	-0,034	-0,045	0,105	0,012	-0,016	-0,043	-0,073	-0,039	-0,124	0,268	0,524	0,556	0,481
3	0,017	-0,024	1,000	0,328	0,309	0,279	0,216	0,041	-0,050	0,074	0,318	0,320	0,323	0,915	0,053	0,000	0,000	0,000	0,000	0,792	0,766	0,438	0,103	-0,688	0,008	0,061	0,030	-0,027	-0,100
4	0,836	0,733	0,328	1,000	0,955	0,850	0,633	0,464	-0,025	-0,206	0,280	0,113	0,118	0,328	0,001	0,063	-0,041	-0,055	0,010	0,327	0,292	0,152	-0,046	-0,237	0,105	0,418	0,462	0,318	0,142
5	0,746	0,705	0,309	0,955	1,000	0,965	0,811	0,311	-0,029	-0,164	0,313	0,114	0,123	0,305	0,020	0,047	-0,031	-0,055	-0,008	0,314	0,264	0,128	-0,037	-0,211	-0,015	0,323	0,471	0,355	0,138
6	0,621	0,620	0,279	0,850	0,965	1,000	0,930	0,192	-0,025	-0,121	0,311	0,111	0,119	0,273	0,045	0,026	-0,025	-0,049	-0,022	0,285	0,228	0,111	-0,025	-0,179	-0,072	0,221	0,438	0,348	0,114
7	0,414	0,437	0,216	0,633	0,811	0,930	1,000	0,076	-0,015	-0,068	0,265	0,092	0,100	0,211	0,083	-0,004	-0,021	-0,035	-0,034	0,220	0,162	0,091	-0,004	-0,126	-0,087	0,087	0,355	0,275	0,055
8	0,462	0,199	0,041	0,464	0,311	0,192	0,076	1,000	0,052	-0,172	-0,390	-0,069	-0,003	0,045	-0,035	0,036	-0,028	-0,014	0,029	0,022	0,038	0,063	-0,081	-0,032	0,722	0,178	0,063	0,004	-0,015
9	-0,020	-0,025	-0,050	-0,025	-0,029	-0,025	-0,015	0,052	1,000	-0,022	-0,051	-0,038	-0,023	0,000	-0,020	-0,020	-0,020	-0,020	-0,020	-0,027	0,021	-0,058	-0,014	0,016	0,038	-0,028	-0,017	-0,009	-0,005
10	-0,230	-0,154	0,074	-0,206	-0,164	-0,121	-0,068	-0,172	-0,022	1,000	0,020	0,019	-0,003	0,086	0,106	-0,026	-0,059	0,006	-0,059	0,040	0,032	0,074	-0,006	-0,131	-0,165	-0,124	-0,076	-0,041	-0,022
11	0,111	0,224	0,318	0,280	0,313	0,311	0,265	-0,390	-0,051	0,020	1,000	0,394	0,156	0,297	0,040	0,000	-0,049	-0,116	-0,027	0,403	0,347	0,177	0,070	-0,195	-0,288	0,011	0,150	0,178	0,110
12	-0,044	-0,007	0,320	0,113	0,114	0,111	0,092	-0,069	-0,038	0,019	0,394	1,000	0,100	0,301	-0,077	-0,014	-0,002	-0,134	-0,011	0,667	0,389	0,181	0,021	-0,211	0,077	-0,177	-0,002	0,055	0,046
13	-0,020	-0,030	0,323	0,118	0,123	0,119	0,100	-0,003	-0,023	-0,003	0,156	0,100	1,000	0,304	-0,053	0,073	-0,002	-0,052	-0,119	0,289	0,266	0,440	0,030	-0,227	-0,021	0,020	-0,007	0,001	-0,063
14	0,024	-0,018	0,915	0,328	0,305	0,273	0,211	0,045	0,000	0,086	0,297	0,301	0,304	1,000	0,113	-0,033	-0,023	-0,013	-0,003	0,783	0,765	0,383	0,074	-0,828	0,010	0,070	0,022	-0,018	-0,097
15	-0,006	-0,003	0,053	0,001	0,020	0,045	0,083	-0,035	-0,020	0,106	0,040	-0,077	-0,053	0,113	1,000	-0,084	0,020	0,260	0,054	0,014	-0,008	-0,050	-0,014	-0,163	-0,022	0,020	-0,030	0,033	-0,020
16	0,083	0,113	0,000	0,063	0,047	0,026	-0,004	0,036	-0,020	-0,026	0,000	-0,014	0,073	-0,033	-0,084	1,000	-0,085	-0,085	-0,085	-0,008	0,007	-0,009	-0,094	-0,018	-0,007	-0,037	0,087	0,032	0,110
17	-0,031	-0,034	0,000	-0,041	-0,031	-0,025	-0,021	-0,028	-0,020	-0,059	-0,049	-0,002	-0,002	-0,023	0,020	-0,085	1,000	-0,085	-0,085	-0,029	-0,070	0,034	0,056	-0,036	-0,007	-0,010	0,008	-0,039	-0,020
18	-0,031	-0,045	0,000	-0,055	-0,055	-0,049	-0,035	-0,014	-0,020	0,006	-0,116	-0,134	-0,052	-0,013	0,260	-0,085	-0,085	1,000	-0,085	-0,052	-0,062	-0,062	-0,069	-0,046	0,014	0,016	-0,032	-0,039	-0,020
19	0,075	0,105	0,000	0,010	-0,008	-0,022	-0,034	0,029	-0,020	-0,059	-0,027	-0,011	-0,119	-0,003	0,054	-0,085	-0,085	-0,085	1,000	0,004	0,029	0,023	-0,014	-0,030	-0,007	0,016	-0,032	0,102	0,110
20	0,030	0,012	0,792	0,327	0,314	0,285	0,220	0,022	-0,027	0,040	0,403	0,667	0,289	0,783	0,014	-0,008	-0,029	-0,052	0,004	1,000	0,710	0,375	0,058	-0,616	0,057	-0,040	0,071	0,000	-0,043
21	0,017	-0,016	0,766	0,292	0,264	0,228	0,162	0,038	0,021	0,032	0,347	0,389	0,266	0,765	-0,008	0,007	-0,070	-0,062	0,029	0,710	1,000	0,371	0,038	-0,570	0,040	0,060	-0,033	-0,005	-0,044
22	0,004	-0,043	0,438	0,152	0,128	0,111	0,091	0,063	-0,058	0,074	0,177	0,181	0,440	0,383	-0,050	-0,009	0,034	-0,062	0,023	0,375	0,371	1,000	0,263	-0,274	0,071	0,010	0,036	-0,060	-0,080
23	-0,078	-0,073	0,103	-0,046	-0,037	-0,025	-0,004	-0,081	-0,014	-0,006	0,070	0,021	0,030	0,074	-0,014	-0,094	0,056	-0,069	-0,014	0,058	0,038	0,263	1,000	-0,061	-0,041	-0,025	0,000	-0,066	-0,033
24	-0,058	-0,039	-0,688	-0,237	-0,211	-0,179	-0,126	-0,032	0,016	-0,131	-0,195	-0,211	-0,227	-0,826	-0,163	-0,018	-0,036	-0,046	-0,030	-0,616	-0,570	-0,274	-0,061	1,000	0,008	-0,040	-0,053	-0,044	0,068
25	0,116	-0,124	0,008	0,105	-0,015	-0,072	-0,087	0,722	0,038	-0,165	-0,288	0,077	-0,021	0,010	-0,022	-0,007	-0,007	0,014	-0,007	0,057	0,040	0,071	-0,041	0,008	1,000	-0,236	-0,145	-0,078	-0,042
26	0,462	0,268	0,061	0,418	0,323	0,221	0,087	0,178	-0,028	-0,124	0,011	-0,177	0,020	0,070	0,020	-0,037	-0,010	0,016	0,016	-0,040	0,060	0,010	-0,025	-0,040	-0,236	1,000	-0,097	-0,053	-0,028
27	0,519	0,524	0,030	0,462	0,471	0,438	0,355	0,063	-0,017	-0,076	0,150	-0,002	-0,007	0,022	-0,030	0,087	0,008	-0,032	-0,032	0,071	-0,033	0,036	0,000	-0,053	-0,145	-0,097	0,000	-0,032	-0,017
28	0,408	0,556	-0,027	0,318	0,355	0,348	0,275	0,004	-0,009	-0,041	0,178	0,055	0,001	-0,018	0,033	0,032	-0,039	-0,039	0,102	0,000	-0,005	-0,060	-0,066	-0,044	-0,078	-0,053	-0,032	1,000	-0,009
29	0,284	0,481	-0,100	0,142	0,138	0,114	0,055	-0,015	-0,005	-0,022	0,110	0,046	-0,063	-0,097	-0,020	0,110	-0,020	-0,020	0,110	-0,043	-0,044	-0,080	-0,033	0,068	-0,042	-0,028	-0,017	-0,009	1,000

Taulukko 10: Regressiomalleissa käytettyjen selittäjien korrelaatio

Malli	1-1	2-1	3-1	5-1	10-1	1-2	2-2	3-2	5-2	10-2	1-1M	1-2M
Kuntoisuus	2 028,6	38,9	33,7	3,3	1,8	31,6	34,4	33,7	3,3	3,1	15,6	32,5
Kertaluokka	1	52	60	615	1127	64	59	60	615	654	130	62

Taulukko 11: Regressiomallien kuntoisuusasteet

7.3.2.3 Residuaalien homoskedastisuus, normaalius ja autokorrelaatio

Homoskedastisuus

Kuvissa 8 ja 9 on esitetty muodostettujen regressiomallien residuaalit sekä niiden jakauma. Residuaalit on tässä määritetty koko havaintoaineistosta (n=408).

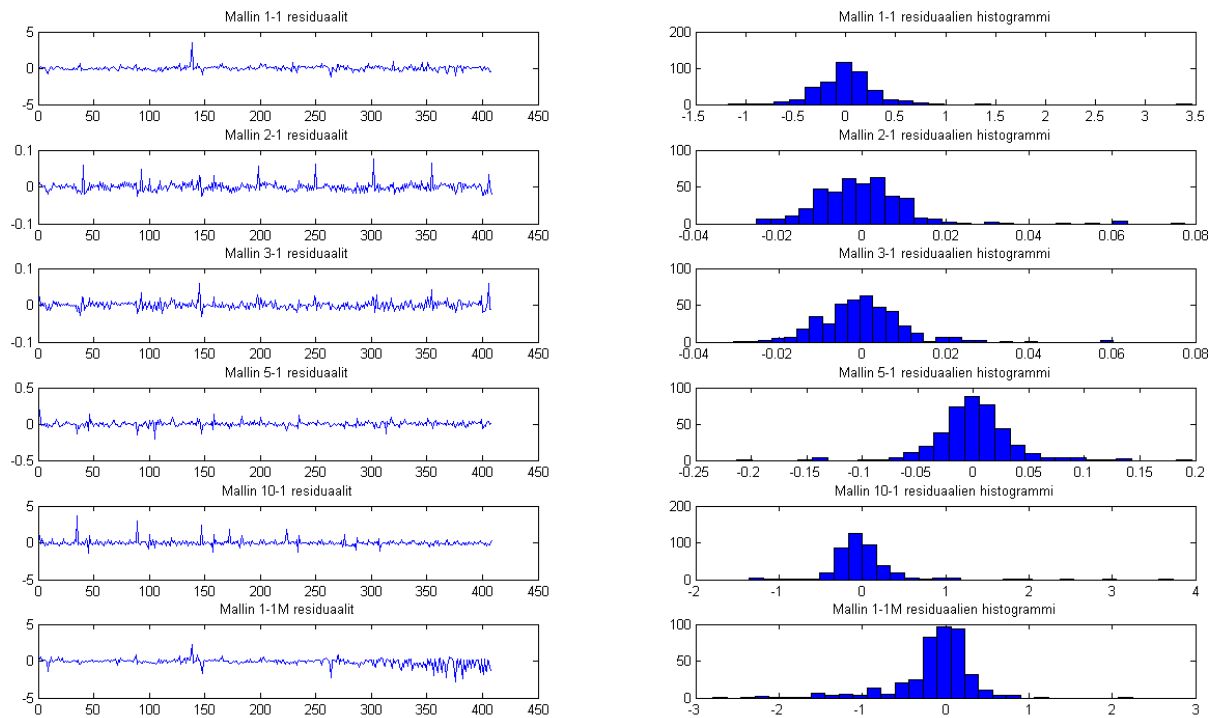
Regressiomallien residuaaleista voidaan huomata seuraavaa: mallien 1-1 ja 1-2 (ker-
tapelit) residuaalit vaikuttaisi hieman heteroskedastisilta. Tämä voidaan huomata
siitä, että residuaalien vyö vaikuttaisi hieman levenevän datassa eteenpäin (oikealle)
siirryttäessä. Tämä on luultavasti seurausta siitä, että ajan myötä mallit yliennusta-

vat potien vaikutusta Loton myyntiin. 2-5 viikon kestoilla vaikuttaisi siltä, että residuaalit ovat suhteellisen homoskedastisia, sillä merkittävää muutosta ei ole residuaalien vyössä havaittavissa. 10 viikon kestolla on taas puolestaan heteroskedastisuutta havaittavissa. Residuaalien vyö selkeästi levenee datan loppua kohden. Mallit 1-1M ja 1-2M vaikuttavat olevan selkeästi heteroskedastisia. Residuaalikuvaajista voidaan huomata, että datan loppuvaiheessa mallit ovat selkeästi yliennustaneet kertapeliä myyntiä.

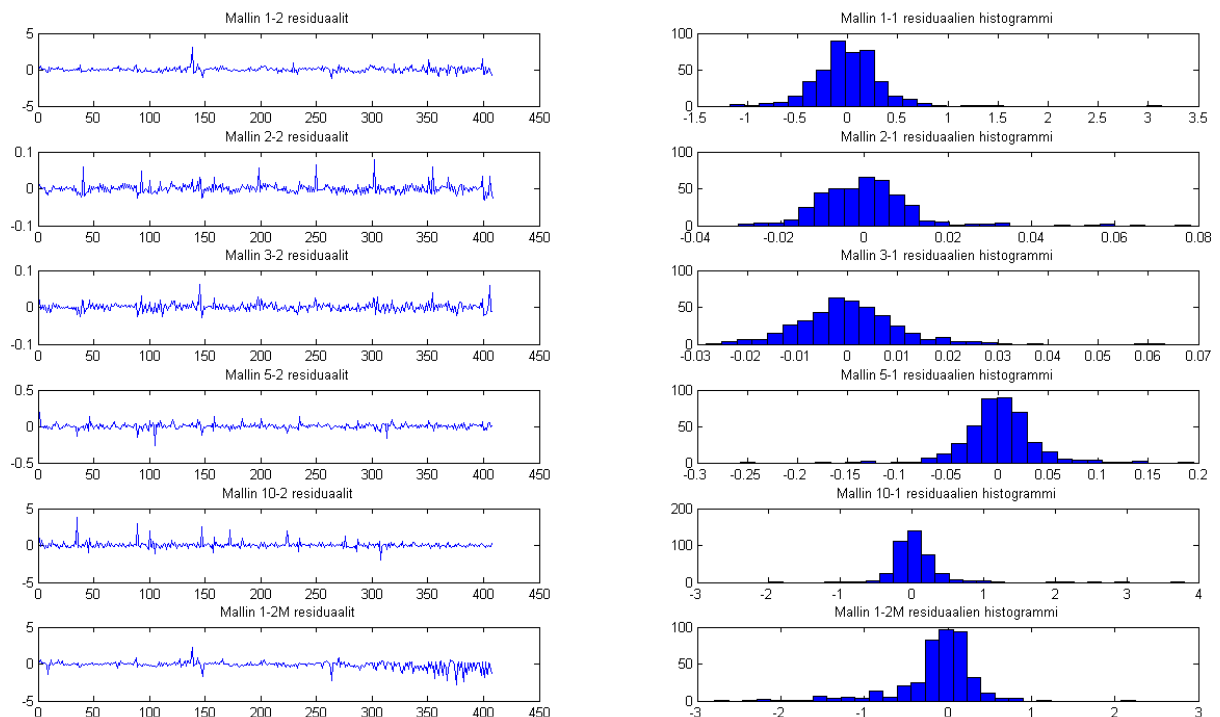
Malleissa on siis havaittavissa lievää heteroskedastisuutta kertapeliä ja kymmenen viikon kestojen osalta. Tämän vahvistaa myös tilastollinen testaus, joka suoritettiin Koenkerin muunnelmalla (Koenker, 1981) Breusch-Pagan heteroskedastisuustestistä. Testin nojalla juuri kertapeliä ja kymmenen viikon kestojen myyntien regressiomallien homoskedastisuus oletuksen ei voi katsoa pätevän. Myös mallit 1-1M ja 1-2M osoittautuvat lievästi heteroskedastisiksi, vaikka kuvaajien perusteella olisi voinut olettaa heteroskedastisuuden olevan todella selvää.

Heteroskedastisuus vaikuttaa lähinnä regressiokertoimien tulkintaan, koska niiden luottamusväleistä tulee tarpeettoman suuria (Mellin; Lineaarinen regressioanalyysi). Simuloinnissa tämän ongelman voidaan katsoa olevan pieni, koska regressiokertoimien tulkinnasta ei simulointivaiheessa olla kiinnostuneita. Toki on tiedostettava, että mallit saattavat hieman yliennustaa Loton myyntiä suuremmilla pottitasoilla. Mallien 1-1M ja 1-2M osalta voidaan todeta, että heteroskedastisuus vaikuttaa regressiokertoimien tulkintaan Loton potin vaikutuksen osalta. Uskoisin kuitenkin, että vaikka heteroskedastisuus näissä malleissa pieni ongelma onkin, pystytään niiden avulla suhteellisen luotettavasti tarkastelemaan sitä, mitkä tekijät Loton kertapelaamiseen vaikuttavat.

Heteroskedastisuus kertapeliä mallien osalta johtunee ainakin osittain siitä, että kierroksella 6/2008 päävoittoluokan osuutta kasvatettiin edellä mainitusti 10 % → 13,6 %. Vaikka Loton pelaaminen kasvoi päävoittoluokan osuuden muutoksen myötä, niin mallien mukaan kasvun olisi voinut olettaa olevan vielä suurempaa. Heteroskedastisuus ei ainakaan silmämääräisesti näyttäisi olevan ongelma malleilla 1-1 ja 1-2, joiden voidaan katsoa suhteellisen hyvin selviytyvän kasvaneen päävoittoluokan vaikutuksista.



Kuva 8: Regressiomallin 1 residuaalit ja niiden jakauma



Kuva 9: Regressiomallin 2 residuaalit ja niiden jakauma

Residuaalien normaalius

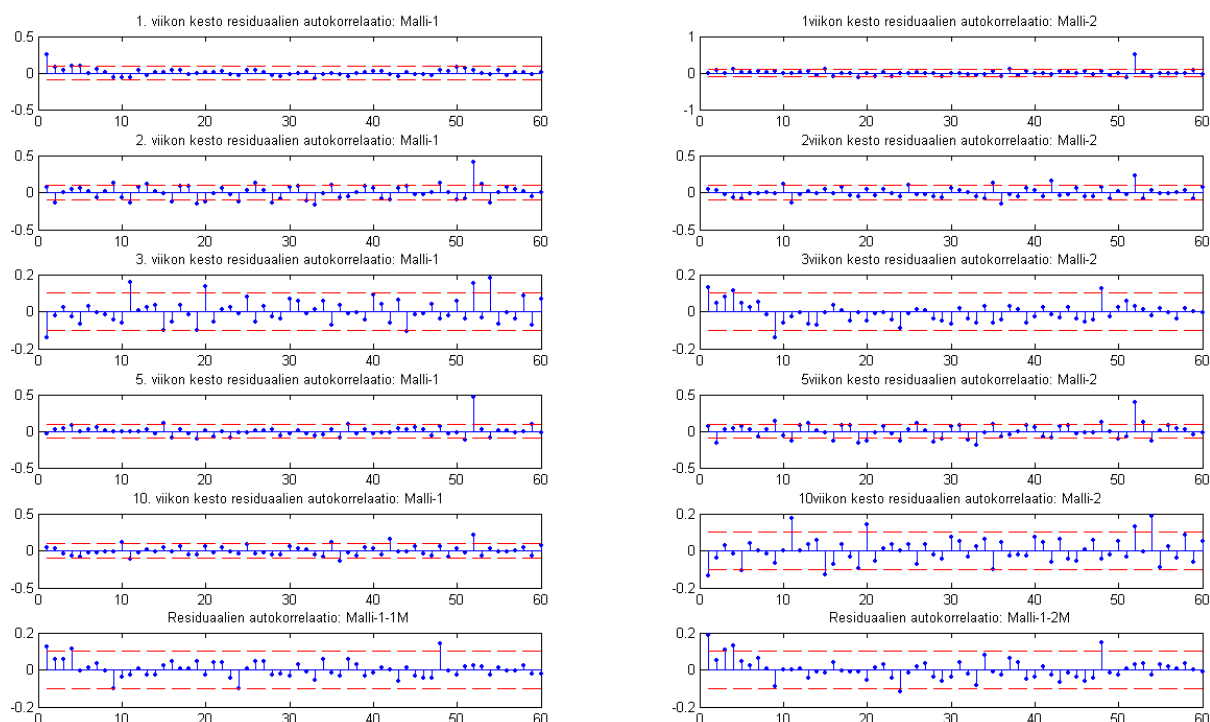
Kuvissa 8 & 9 on esitetty regressiomallien residuaalien jakauma histogrammina. Silmämaisesta vaikuttaisi siltä, että residuaalit voisivat hyvin ollakin normaalijakautuneita (tiedostaen että joukossa on hyvinkin suuria residuaalien arvoja). Tilastollinen

testaus residuaalien normaaliudesta (Jarque, Bera; 1987) kuitenkin kumoaa tämän oletuksen kaikkien regressiomallien osalta. Näin ollen tilastollisen päättelyn t- ja F-jakaumiin nojautuen ei välttämättä voida katsoa olevan enää pätevää (Mellin; Lineaarinen regressioanalyysi). Otoskoko on tässä kuitenkin kohtalaisen suuri ($N=408$), joten voidaan olettaa, että muodostettujen regressiomallien avulla voidaan suorittaa suuntaa antavaa tilastollista päättelyä Loton myyntiin vaikuttavista tekijöistä (Mellin; Lineaarinen regressioanalyysi). Tässä ei sen tarkemmin tutkita syitä, miksi residuaalit eivät normaaliustestiä läpäise. Yhtenä vaihtoehtona on se, että aika ajoin Lotossa on ollut erittäin korkeita tai alhaisia myyntejä, joita mallit eivät ole kyenneet selittämään. Nämä merkittävästi poikkeavat havainnot lienevät merkittävässä roolissa normaaliuden hylkäämisessä.

Residuaalien autokorrelaatio

Kuvassa 9 on esitetty regressiomallien autokorrelaatiot, sekä niille määritetty raja, jolla autokorrelaation voidaan katsoa olevan tilastollisesti merkittävä. Kuvasta voidaan huomata, että pääsääntöisesti autokorrelaatiot sijoittuvat raja-arvojen väliin. 95 prosentin luottamustaso on tässä määritetty kaavalla $\pm \frac{2}{\sqrt{N}}$, missä $N = 408$ kuvaa havaintojen lukumäärää (Pindyck et. al.; 1991).

Osan autokorrelaatioista voidaan havaita olevan merkittäviä. Merkittäväksi nousevat autokorrelaatiot ovat usein joko yhden tai sitten 52 havainnon välein. Tulosta voidaan pitää luonnollisena, sillä esimerkiksi vuosisyklin voidaan olettaa vaikuttavan Loton pelaamiseen. Näin ollen joidenkin selitettävien muuttujien kohdalla differensointi olisi saattanut olla perusteltua. Differensointia ei kuitenkaan toteutettu, sillä tulosten katsottiin olevan riittävän hyviä simulointeja varten muodostetuilla malleilla. Residuaalien autokorrelaatio vaikuttaa lähinnä regressiokertoimien luottamustasoihin ja niiden testisuureiden arvoihin. Simulaatioiden kannalta tällaisten vaikutusten voidaan katsoa olevan vähemmän merkityksellisiä.



Kuva 10: Regressiomallien residuaalien autokorrelaatio

7.3.2.4 Regressiomallien ennustuskyky ja sen tilastollinen testaus

Regressiomallien ennustuskyvyn graafinen tarkastelu

Kuvissa 11 ja 12 on kuvaajien avulla esitetty estimoitujen regressiomallien ennustuskykyä. Verrokkidatana on ollut vuoden 2010 toteutuneet Loton myynnit. Regressiomallien estimointidata päättyi vuoden 2009 loppuun.

Kuvassa 11 on esitettyinä mallien 1-1 ja 1-2, sekä mallien 1-1M ja 1-2M ennusteet vs. toteutuneet kertapeliin myynnit. Kuvassa 12 on vastaavasti esitetty kestojen ennusteet vuoden 2010 toteumiin verrattuna. Kuvissa on käytetty seuraavaa notaatiota:

- musta viiva ja ympyrä: toteutunut myynti
- Punainen viiva ja risti: mallien 1 ja 2 estimaattorit
- Sininen viiva ja timantti: mallit 1-1M ja 1-2M

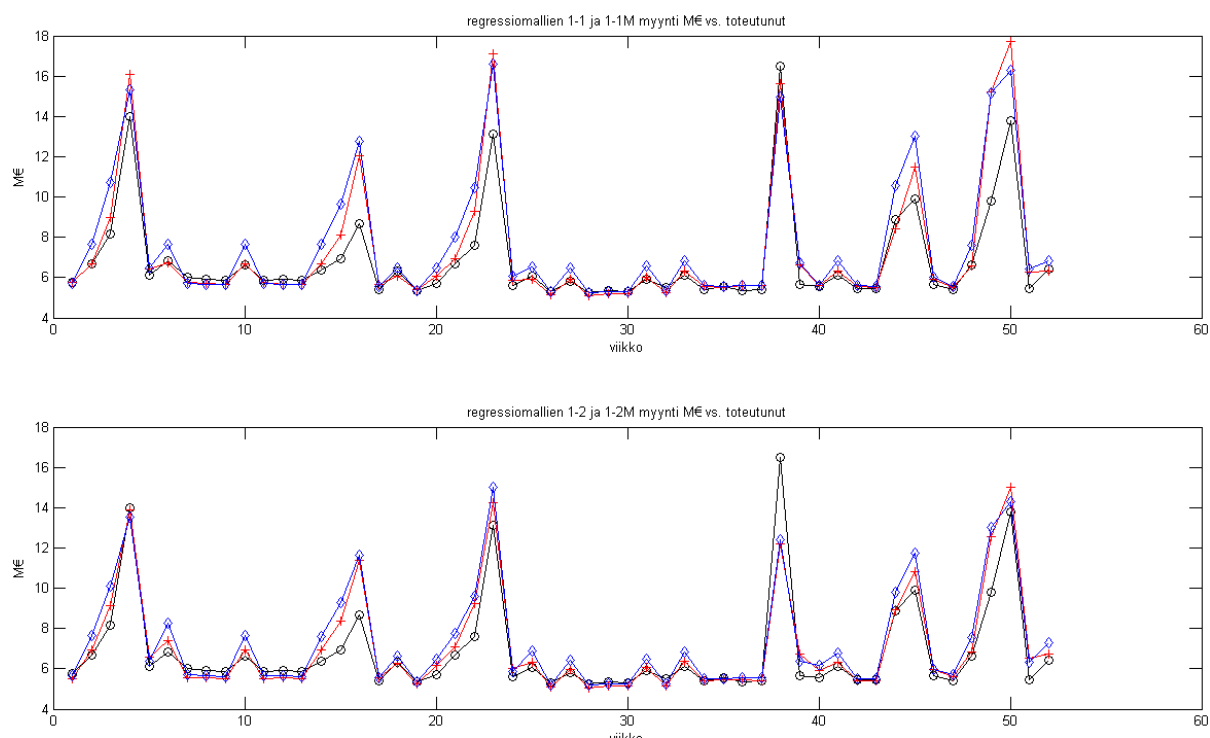
Kuvasta 11 voidaan havaita, että muodostetut regressiomallit ennustavat kertapeliin myyntiä varsin hyvin, etenkin alhaisemmilla myyntitasoilla. Suurempia myyntejä mallit useammin yliarvioivat. Tämä yliarviointi näyttäisi jonkin verran korostuvan malleilla 1-1 ja 1-1M. Yliarvioinnin voidaan ainakin osittain olettaa johtuvan jo edellä kuvatuista pottien inflatoitumisesta, mikä selittynee valtaosin edellä kuvatulla päävoitoluokan osuuden muutoksella. Kierroksen 38/2010 voidaan havaita olevan poikkeuksellinen siinä mielessä, että tuolloin kertapeliin myynti ylitti mallien antamat ennusteet. Kierroksella 38/2010 Lotossa oli Veikkauksen 70-vuotisjuhlallisuuksien

kunniaksi juhlapotti 7,2 miljoonaa euroa (tuohon mennessä Loton historian suurin potti), mikä sai ihmiset pelaamaan todella innokkaasti.

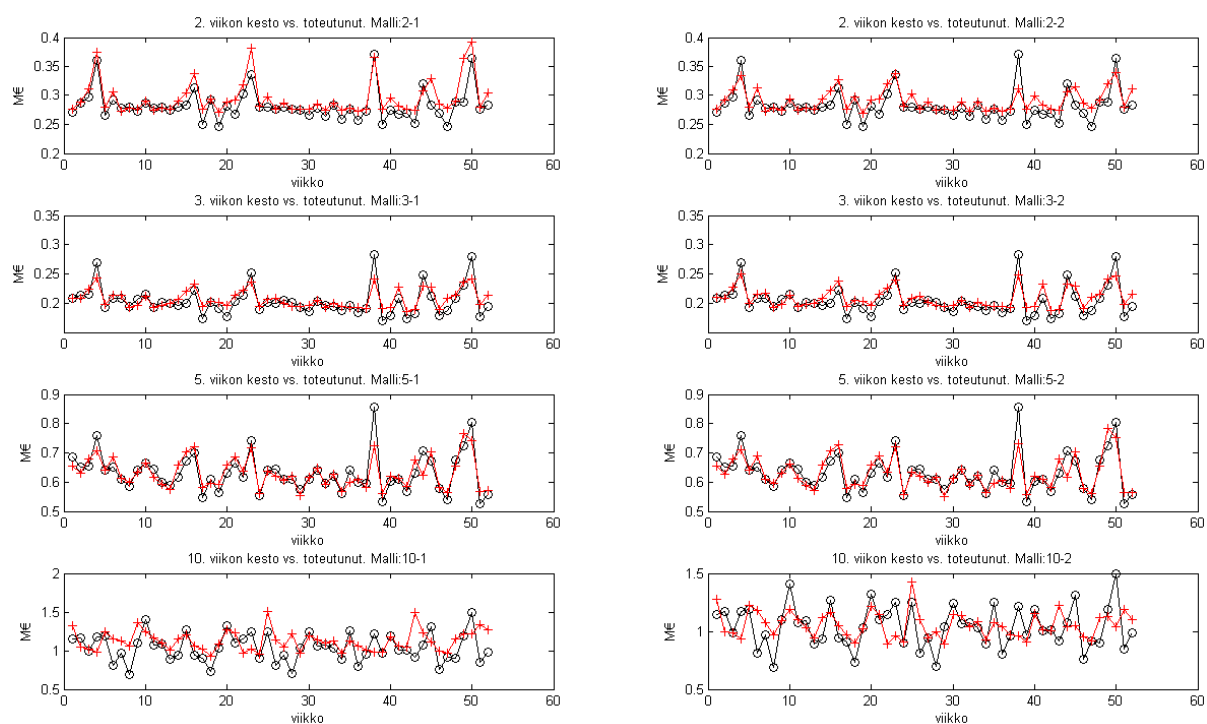
Kuvasta 12 voidaan havaita, että mallit kykenevät ennustamaan myös kestopelaamista suhteellisen hyvin. 10 viikon kestopelaamisen ennustamisen havaitaan kuitenkin olevan ainoastaan keskinkertaista, trendit siitäkin on kuitenkin saatu poimittua. Lyhyiden kestopelien mallien kohtalaisen hyvä suoriutuminen ja 10 viikon kestojen ennustamisen haasteet olivat odotettavissa mallien selityksasteiden perusteella.

Kahden viikon kestopelien osalta voidaan havaita mallien ennusteiden lievää yliarvioimista vuoden 2010 loppupuolella. Mallien 2-1 ja 2-2 ennusteet ovat suurempia kuin toteutuneet myynnit lähes järjestään, mitä lähemmäksi loppuvuotta 2010 mennään. Kolmen viikon kestoilla vastaavaa ilmiötä on myös varauksin havaittavissa, tosin selkeästi heikompana. Pidemmällä kestoilla vastaavaa ilmiötä ei kuvaajista voi selkeästi nähdä. Lyhyemmällä kestoilla voidaan olettaa, että mallien myynnin yliennustaminen johtuu pääsääntöisesti siitä, etteivät mallit ole kyenneet ottamaan huomioon Loton pottien osittaista inflatoitumista.

Kestopelien regressiomalleista voidaan yleisesti havaita, että mallit harvemmin merkittävästi yliarvioivat pelaamisen määrää. Tämä on täysin päinvastainen tulos kuin kertapelien osalta. Tämä selittyy sillä, että kestopelaamisen voidaan havaita olevan selvästi stabiilimpaa kuin kertapelaamisen. Siihen vaikuttavat ulkoiset tekijät, kuten Loton potti, selkeästi vähemmän kuin esimerkiksi kertapelaamiseen. Kestopelaamiseen vaikuttavat voimakkaasti tietyllä tavalla kertaluontoiset tekijät, eli esimerkiksi hinnankorotukset sekä mahdolliset 10+1 kampanjat.



Kuva 11: Kertapeli regressiomallien ennustuskyky (1/2010-52/2010)



Kuva 12: Kestopeli regressiomallien ennustuskyky (1/2010-17/2010)

Regressiomallien ennustuskyvyn tilastollinen testaus

Regressiomallien graafisen tarkastelun lisäksi malleille suoritettiin ennustuskkyä mittaava tilastollinen testaus. Hyödynnettävänä tilastollisen testinä oli χ^2 - testisuure (Mellin; Lineaarinen regressioanalyysi), joka määritettiin kullekin regressiomallille seuraavasti:

$$\chi^2 = \sum_{i=N+1}^{N+h} \frac{u_i^2}{s^2} \quad (18)$$

, missä N kuvaa regressioanalyysissä käytettyjen havaintojen lukumäärää ($N = 408$), h kuvaa verrokkihavaintojen lukumäärää ($h = 52$), u_i kuvaa ennustevirhettä havainnolla i ja s^2 on harhaton estimaattori tarkasteltavan regressiomallin jäännösvarianssille.

Taulukossa 12 on esitetty kaavassa (18) määritetyn testisuureen arvot kullekin regressiomallille. Taulukkoon on myös lisätty viiden prosentin merkitsevyystason raja-arvo, jolla mallien regressiokertoimien vakioisuus oletus tulisi hylätä.

Taulukosta voidaan huomata, että ainoastaan viiden ja kymmenen viikon kestot läpäisevät tilastollisen testin. Tulokset ovat selkeästi heikommat, mitä parametrien vakioisuutta testattaessa. Tämä selittyy pääosin sillä, että vuonna 2010 oli voimassa Loton korkeampi päävoittoluokan osuus: 13,6%, kun valtaosassa estimointitaidaa päävoittoluokan osuus oli 10%. Näin ollen muodostetut mallit helposti yliarvioivat myyntejä johtuen Loton pottien kasvusta.

Kertapeliä osalta voidaan havaita, että mallit ajoittain selkeästi yliarvioivat myyntiä. Voidaan osoittaa, että itse asiassa nämä yksittäiset ylilyönnit pääsääntöisesti selittävät kertapeliä todella suuret χ^2 - testisuureiden arvot. Kahden ja kolmen viikon kestojen mallien osalta lievä yliennustaminen on myös osana selittämässä suurehkoja χ^2 - testisuureen arvoja.

Tarkastellessa mallien absoluuttista ennustuskkyä voidaan niiden suoriutumista pitää kuitenkin riittävän hyvänä, jotta mallien avulla voidaan vuositasoon myyntejä Lotolle simuloida. Simuloinnin tärkeimpiä tavoitteita on vertailla keskenään erilaisia pottitarjoamia sekä tarkastella sitä, kuinka paljon vuositasolla Loton myynnit vaihtelevat tarjolla olevista poteista riippuen. Näiden simulointitavoitteiden osalta voidaan todeta, ettei lievä myyntien yliarviointi suuremmilla poteilla ratkaisevasti vaikuta simulointituloksiin. Loton myyntien mallintamisen ollessa kyseessä on vain hyväksyttävä että myynteihin vaikuttaa moni tekijä, joita on hyvin vaikea mallien sisään integroida.

Malli	1-1	2-1	3-1	5-1	10-1	1-2	2-2	3-2	5-2	10-2	1-1M	1-2M
Chi-Square	781,8	137,8	88,0	40,0	13,0	377,6	109,0	87,8	37,3	8,6	721,3	302,7
Raja-arvo	70	70	70	70	70	70	70	70	70	70	70	70

Taulukko 12: Mallien ennustekyvyn tilastollinen testaus

7.4 LOTON MYYNIN MALLINTAMISEN HAASTEITA JA JATKOKEHITYS MAHDOLLISUUKSIA

Loton myynnin mallintaminen oli erittäin haasteellinen tehtävä. Heti alkuvaiheessa tuli selväksi, että vaikka voitiin selkeästi identifioida myyntiin vaikuttavia tekijöitä, paljon myyntiin vaikuttavia tekijöitä jäisi mallien ulkopuolelle. Syitä tähän on monia: osa myyntiin vaikuttavista tekijöistä on äärimmäisen hankala regressiomalleihin sisällyttää, vaikka ne voitaisiin identifioida. Osa myyntiin vaikuttavista tekijöistä ei keksitty kuin vasta jälkikäteen.

Säätilan on havaittu selkeästi vaikuttavan Loton pelaamiseen. Säätilan vaikutuksesta voisi karrikoidusti sanoa, että äärimmäiset sääolosuhteet laskevat Loton pelaamista. Äärimmäisillä olosuhteilla tässä tarkoitetaan esimerkiksi erittäin kovaa pakkasta, runsaita vesisateita tai vaikka erittäin kaunista kesäkeliä. Säätilaa ei kuitenkaan ole sisällytetty malleihin millään tavalla. Tulevaisuudessa vastaavia mallinnoksia tehtäessä olisikin mielenkiintoista pyrkiä sisällyttämään malleihin tieto siitä, millainen säätila on pelikierroksella vallinnut.

Säätilan lisäksi näihin regressiomalleihin ei sisällytetty markkinointitoimenpiteitä. Muissa työtehtävissä olen tutkinut markkinointitoimenpiteiden vaikutusta Loton pelaamiseen ja voin todeta, että myös niiden selityskykyä näissä malleissa olisi voitu hyödyntää. Tulevaisuuden kehitystyöksi jääkin myös se, että markkinointitoimenpiteet ovat mukana ja ne sisällytetään myös simulaattorin toimintaan.

Kuten diagnostisen testauksen aikana tuli ilmi, on selkeästi havaittavissa, että myös Lottoon ja erityisesti Loton pottiin kohdistuu inflatoitumista. Inflatoituminen korostuu etenkin ajankohtina, jolloin Loton potteja on kasvatettu päävoiton osuutta nostamalla. Tätä Loton pottien inflatoitumista en kyennyt näihin regressiomalleihin sisällyttämään. Tulevaisuuden tehtäväksi jääkin selvittää, voidaanko tätä inflatoitumista kuvata paremmin joillain muilla matemaattisilla menetelmillä tai mahdollisesti jaloitamalla tässä työssä hyödynnettyjä menetelmiä.

Työn edetessä tuli myös ilmi, että joitain potentiaalisia selittäjäkandidaatteja jäi kokonaan testaamatta. Eräs potentiaalinen selittäjä olisi ollut I/O muuttuja, joka olisi indikoinut sen, onko tarkasteluviikolla ollut tarjolla Loton historian suurin jättipotti. Tämän on huomattu vaikuttavan myyn-teihin ajoittain hyvinkin voimakkaasti (esimerkiksi kierros 38/2010) ja tällä muuttujalla olisi ehkä voitu mallintaa Loton myynnin ”epärationalista” käyttäytymistä ennätysuurten pottien aikaan. Toinen mahdollisesti hyvinkin toimiva selittäjä 10 viikon kestopelaamiseen olisi ollut I/O muuttuja,

joka olisi indikoinut sitä, että kyseisellä kierroksella on ollut pelaajille tarjolla 10+1 kampanja.

8 LOTON PELIMYYNNIN SIMULOINTI

8.1 SIMULOINTIMALLIN KUVAAUS

Kuvassa 13 on esitetty periaatekuva luodun simulaattorin toimintaperiaatteesta yksittäisen vuoden myyntien sekä voittorahaston kulutuksen määrittämisen osalta. Lähtödatana simuloinnissa oli vuoden 2009 Loton 10 viimeisen kierroksen regressiomalleja vastaava selittäjädata.

Simulaattorin toimintaperiaate voidaan kuvata seuraavasti. Kunakin vuonna on simulaatioissa ollut yhteensä 52 arvontaa (kierrosta). Vuositason simulointeja suoritettiin puolestaan yhteensä 50 000 kappaletta. Kunkin kierroksen osalta määritettiin ensimmäiseksi regressiomallien avulla kierroksen myyntiestimaatit kestoittain. Tämän jälkeen määritettiin kierrosvaihdon estimaatti kaavan (6) periaatteen mukaisesti. Kierrosvaihtoestimaatin nojalla laskettiin kierrokselle osallistuneiden pelirivien lukumäärä ja edelleen arvio siitä, kuinka suuri osa Loton riviavaruudesta on kierroksella pelattu eli ruudukoiden täyttöaste. Ruudukoiden täyttöasteen määrittämisen menetelmä on kuvattu tarkemmin tämän kappaleen loppupuolella. Seuraavaksi suoritettiin satunnaisluvun määrittäminen väliltä 0-1 ja verrattiin sitä edellä määritettyyn täyttöasteeseen. Vertailun nojalla määritettiin, löytyikö tarkasteltavalla kierroksella Loton päävoittoa vai ei. Mikäli päävoittoa ei löytynyt, siirryttiin pottityypissä yksi askel eteenpäin. Mikäli päävoitto puolestaan löytyi, asetettiin seuraavaksi Loton kierrokseksi peruskierros (pottityyppi 0). Simuloinneissa suurimmaksi pottityypiksi asetettiin 5, mikä tarkoittaa sitä, että pottityypin 5 tilanteessa arvonnassa löytyi aina päävoitto. Tätä perustellaan sillä, että regressiomalleja muodostettaessa ei ollut käytössä dataa siitä, kuinka Loton pelaaminen kehittyy kun pottityyppi on korkeampi kuin 5. Todellisuudessa myös arvoa 5 suuremmat pottityypit ovat hyvin harvinaisia, ajanjaksona 2002-2009 niitä ei esiintynyt lainkaan (ks. taulukko 7). Mikäli arvonnassa löytyi päävoitto, määritettiin mahdollisen voittorahaston tarve potin suuruuden kattamiseen. Voittorahastoa tarvittiin potin kattamiseen, mikäli kyseisen pottijakson yhteenlasketusta vaihdosta määritetty ensimmäisen voittoluokan osuus jäi alle sääntöjenmukaisen ensimmäisen voittoluokan osuuden (13,6 % pottijakson kokonaisvaihdosta). Pottijaksolla tarkoitetaan tässä yhteydessä tarkasteltavan kierroksen ja edellisen kierroksen, jolloin päävoitto löytyi jälkeisen kierroksen välisiä Loton kierroksia.

Edellä kuvattua menetelmää sovellettiin kaikkiin tarkastelun kohteina olleisiin pottitarjoamiin. Tutkinnan kohteena olleet pottitarjoamat on listattu taulukossa 14. Simulointituloksista tallennettiin vuositason myyntitiedot sekä voittorahaston kulutustiedot. Simulointitulokset on esitetty tarkemmin taulukossa 13. Eräänä työn tärkeimpä-

nä ja samalla haasteellisimpana tavoitteena oli vertailla eri pottitarjoamien paremmuutta. Tätä varten muodostettiin pisteytysfunktio, jonka avulla kukin pottitarjoama pisteytettiin simulointitulosten perusteella. Pisteytykset suoritettiin erikseen niin tavanomaisen kuin logistisen regression tuloksille. Pisteytysfunktion periaate on esitetty kaavassa (19):

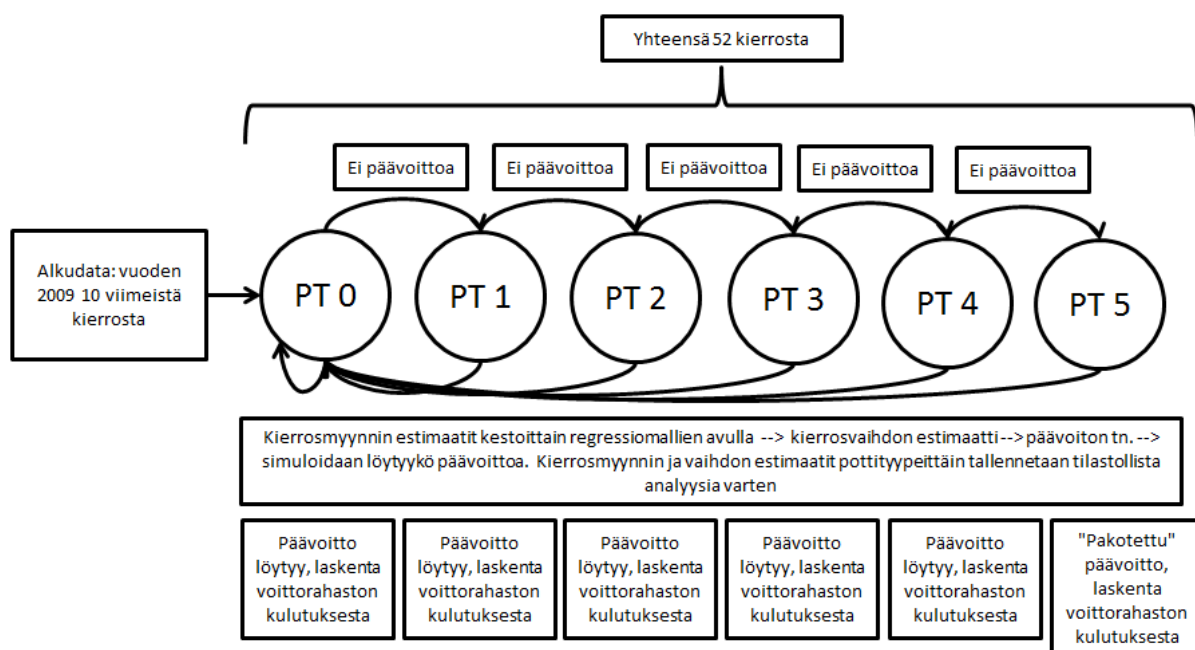
$$\text{Pist}_{ij} = [p_1 \ p_2 \ p_3 \ p_4] * [Mka_{ij} \ Mvar_{ij} \ VRtav_{ij} \ VRvar_{ij}]^T \quad (19)$$

, missä $i = 1 \dots 20$ kuvaa tarkasteltavan pottitarjouksen indeksiä ja $j = 1,2$ kuvaa tarkasteltavaa regressiomallia, 1 = tavallinen ja 2 = logaritminen regressiomalli. p_k kuvaa pisteytyksessä käytettävää painovektoria siten, että: $p_1 = p_2 = p_3 = 0,3$ & $p_4 = 0,1$. Mka tarkoittaa kunkin mallin simulointien perusteella saadusta keskiarvoisesta myynnistä määritettyä pisteytystä. Mka määritettiin siten, että kaikkien pottitarjousten keskimääräistä myyntiä verrattiin keskimääräiseltä myynniltään heikoimman pottitarjousmallin keskimyyntiin. Tämän jälkeen keskimyynti pisteytettiin pottitarjouksittain siten, että paras keskimyynti sai yhden pisteen ja huonoin nolla pistettä. Muut keskimyyntit pisteytettiin lineaarisesti välille 0-1. Vastaavaa logiikkaa käytettiin myös pisteytyksissä $Mvar$, mikä kuvaa pottitarjousmallin myynnin varianssia ja $VRvar$, mikä kuvaa pottitarjousmallin voittorahaston kulutuksen varianssia. Varianssien osalta tulkinta oli se, että mitä alhaisempi varianssi pottitarjousmallilla on sen parempi. $VRtav$ noudattaa muuten samaa logiikkaa, mutta siinä lähtökohdaksi ei asetettu kaikkein alhaisinta voittorahaston kulutusta. Muuttujalla $VRtav$ verrokkina oli se euromäärä, millä kyseinen pottitarjousmalli vuoden aikana voittorahastoa keskimäärin kartuttaa (0,9 % vuosivaihdosta). Tätä voidaan perustella sillä, että mikäli pottitarjousmalli kuluttaisi voittorahastoa suunnilleen sen verran mitä se sitä kartuttaa, olisi peli ikään kuin täysin omavarainen.

Edellä kuvattu pisteytysfunktio muodostettiin, jotta eri pottitarjoukset voitaisiin mahdollisimman objektiivisesti asettaa paremmuusjärjestykseen. Tehtävä oli haastava, mutta mielestäni muodostettu pisteytysfunktio painottaa suhteellisen hyvin Loton dynamiikan kannalta oleellisia asioita. Pisteytysfunktio huomioi sekä vuositasen myynnin odotusarvoa että sen varianssia. Pisteytysfunktio ottaa lukuun myös sen, kuinka omavarainen pottitarjousmalli on voittorahaston kulutuksen suhteen sekä sen, kuinka paljon vuositasolla voittorahaston kulutus vaihtelee.

Pisteytysvektori pohjautuu täysin työnlaatijan näkemykseen ja kokemukseen siitä, mitä tekijöitä Loton vuositasen tuloksen tarkasteluissa tulisi painottaa ja kuinka niiden tulisi suhtautua toisiinsa. Kuten pisteytysfunktioista nähdään, myynnilliset tekijät yhdessä voittorahaston tavoitteellisen käytön lisäksi koettiin kaikkein merkittävimmiksi.

Simulaattori toteutettiin siten, että päävoittoluokan osuus vaihdosta oli 13,6%, kuten se oli Lotossa vuonna 2010 (lähtödatana oli vuoden 2009 loppudata).



Kuva 13: Periaatekuva simulaattorin toimintaperiaatteesta

Ruudukoiden täyttöaste ja sen määrittäminen

Ruudukoiden täyttöasteen kehittyminen rivimäärien mukaan on esitetty liitteissä olevassa kuvassa 20. Täyttöasteella tarkoitetaan tässä sitä, että kuinka suuri osa Loton pelirivien kombinaatioavaruudesta tietyllä kierroksella pelattu. Täyttöasteen H estimaatti on määritetty kaavan 20 mukaan (Järvinen; 2007) ja näin ollen todennäköisyys ettei tietyllä kierroksella löydy yhtään päävoittoa on $1 - H = e^{kN}$.

$$H = 1 - e^{kN} \quad (20)$$

Kaavassa 20 N kuvaa Loton kierrokselle osallistuneiden pelirivien lukumäärää. k puolestaan kuvaa kulmakerrointa, joka voidaan tilastollisten aineistojen avulla Lotolle määrittää. Tilastollisilla aineistoilla tarkoitetaan tässä yhteydessä dataa, joka pitää sisällään tiedon siitä, kuinka monta peliriviä kullekin pelikierrokselle on osallistunut ja edelleen, kuinka suuri osa Loton riviavaruudesta on kyseisellä kierroksella pelattu. Tätä työtä varten määritettiin k :n lukuarvo ja tulokseksi saatiin $-5,8999 \cdot 10^{-8}$. Loton täyttöasteen estimaatti määritettiin edellä kuvatulla menetelmällä, sillä Loton riviavaruus ei täyty binomijakauman oletusten mukaisesti (Järvinen; 2007). Tämä voidaan havaita esimerkiksi kuvan 20 ylemmästä kuvasta, josta käy ilmi se, että todellinen täyttöaste on selvästi teoreettista alempi. Tämä selittyy sillä, että suuri osa Loton pelatuista numeroista kasautuu pienille numeroille. Loton täyttöasteen teoriaan voi perehtyä tarkemmin Harri Järvisen diplomityön avulla.

8.2 SIMULOINTITULOKSET

Taulukossa 13 on esitetty simulointien tulokset kaikille testatuille pottitarjoamille. Simuloinnissa mukana olleet pottitarjoamat ovat kuvattuna taulukossa 14.

Taulukossa 13 esitetyissä simulointituloksissa on eri pottitarjoamat pisteytetty paremmuusjärjestykseen asettamista varten. Pottitarjoamien pisteytys on suoritettu edellä kuvatun kaavan (19) perusteella.

Taulukkoa 13 luetaan seuraavasti:

- Värikoodaus: Kunkin kategorian viisi parasta ja viisi huonointa tarjoamaa on korostettu. Vihreä indikoi viiden parhaan ja punainen viiden huonoimman joukkoon kuulumista
- Mallin ind. kuvaa tarkasteltavan pottitarjoaman indeksiä eli tunnusta
- -1 ja -2 loppuliite kuvaa sitä tarkastellaanko mallilla 1 vai mallilla 2 saatuja simulointituloksia. -1 viittaa tavanomaiseen ja -2 logaritmisen regression simulointituloksiin
- Rank kuvaa sitä, kuinka eri pottitarjoamat rankkautuivat kaavassa 19 esitetyn pisteytyksen mukaan
- Pist kuvaa eri tarjoamamallien saamia pisteitä kaavan 19 mallin mukaan
- KA M: tarjoamien vuositason myyntien keskiarvo
- MED M: tarjoamien vuositason myyntien mediaani
- Pelik.: tarjoamien vuositason pelikatteen keskiarvo ($=0,6 \cdot KA\ M$)
- VAR M: tarjoamien vuositason myyntien varianssi
- M X%: tarjoamien myyntien X-persentiili, esimerkiksi M 25% tarkoittaa myynnin arvoa, jonka alle jää 25% vuosista kyseisellä pottitarjoamalla
- KA VR: tarjoamien vuositason voittorahastokulutuksen keskiarvo
- MED VR: tarjoamien vuositason voittorahastokulutuksen mediaani
- VAR VR: tarjoamien vuositason voittorahastokulutuksen varianssi

Pottitarjoamataulukkoa 14 luetaan seuraavasti:

- Mallin ind. kuvaa tarkasteltavan pottitarjoaman indeksiä eli tunnusta
- Pottityyppi-sarakkeisiin on kuvattuna, minkä suuruinen potti on pelaajille ollut tarjottuna kullakin pottityypillä ja pottitarjoamalla
- Pottien tulo on havainnollistava muuttuja, jonka avulla voi nopeasti vertailla eri pottitarjoamien pottien suuruusluokkia. Suuremman pottien tulon voidaan katsoa antavan indikaatioita siitä, kuinka nopeasti Loton potti kasvaa kussakin pottitarjoamassa

Mallin ind.	Rank-1	Rank-2	Pist-1	Pist-2	KA M-1	KA M-2	MED M-1	MED M-2	Pelik.-1	Pelik.-2	VAR M-1	VAR M-2	M 5%-1
1	2	1	0,676	0,682	465	455	463	452	279	273	519	523	431
2	3	2	0,673	0,677	465	454	463	452	279	273	527	531	431
3	8	3	0,630	0,605	468	459	465	457	281	275	600	647	431
4	9	8	0,619	0,577	473	464	471	461	284	278	662	726	435
5	14	10	0,570	0,549	483	469	482	467	290	281	784	768	441
6	5	4	0,638	0,598	475	466	473	463	285	279	651	702	437
7	13	11	0,582	0,540	485	472	483	469	291	283	784	781	443
8	7	6	0,630	0,587	483	475	481	473	290	285	645	692	445
9	15	13	0,555	0,526	496	484	494	482	298	290	787	781	454
10	16	15	0,547	0,510	496	485	494	483	298	291	796	800	453
11	18	17	0,467	0,446	505	492	503	490	303	295	938	904	457
12	19	18	0,459	0,436	504	492	502	490	303	295	947	919	457
13	1	5	0,677	0,588	485	472	483	470	291	283	604	683	449
14	6	9	0,633	0,559	496	482	493	480	297	289	656	711	458
15	11	12	0,590	0,527	504	488	502	486	303	293	736	757	464
16	12	16	0,584	0,504	505	489	503	487	303	293	746	789	464
17	4	7	0,668	0,584	497	483	494	481	298	290	550	630	463
18	10	14	0,595	0,516	513	495	511	493	308	297	692	733	473
19	17	19	0,498	0,401	539	523	537	521	323	314	830	890	495
20	20	20	0,422	0,381	553	543	551	541	332	326	774	811	510
Mallin ind.	M 5%-2	M 25%-1	M 25%-2	M 75%-1	M 75%-2	M 95%-1	M 95%-2	KA VR-1	KA VR-2	MED VR-1	MED VR-2	VAR VR-1	VAR VR-2
1	421	448	438	479	469	506	496	0,8	1,7	0,8	1,7	0,1	0,3
2	421	448	437	479	469	506	496	1,1	2,1	0,9	1,9	0,5	1,0
3	421	450	440	483	475	512	505	1,0	2,0	0,9	1,9	0,3	0,4
4	424	454	444	490	481	519	512	2,7	4,1	2,7	4,1	0,4	0,5
5	427	462	449	501	487	532	518	3,4	5,6	3,4	5,5	1,5	2,3
6	426	456	446	491	483	521	513	3,5	4,7	3,5	4,7	0,6	0,7
7	430	465	451	503	489	534	521	4,7	6,8	4,7	6,7	1,2	1,6
8	436	464	456	499	492	529	522	9,3	10,4	9,3	10,4	3,1	3,6
9	442	476	464	514	502	546	534	11,8	13,6	11,9	13,7	3,7	3,9
10	442	476	464	514	503	546	535	12,3	14,3	12,3	14,4	2,8	2,5
11	445	483	470	524	511	558	544	13,4	15,0	13,6	15,2	3,1	4,3
12	446	482	470	524	511	558	545	13,6	15,5	13,7	15,7	2,6	2,9
13	433	467	453	501	488	530	519	8,3	9,9	8,3	9,9	2,2	2,9
14	442	477	463	512	499	541	530	13,2	14,9	13,3	15,0	3,7	4,7
15	447	485	469	522	506	553	537	15,2	17,4	15,3	17,5	4,2	4,8
16	447	485	469	522	507	553	539	15,6	18,0	15,7	18,2	3,2	3,1
17	446	479	465	511	499	539	528	16,4	17,7	16,3	17,6	11,1	12,3
18	455	494	476	530	513	559	544	20,4	22,4	20,4	22,4	9,8	10,8
19	477	519	501	558	542	589	575	30,5	32,7	30,7	32,9	13,9	15,5
20	500	533	522	571	561	602	593	42,6	43,7	42,6	43,7	41,0	41,7

Taulukko 13: Pottitarjoamien simulointitulosten vertailu

Mallin ind.	Pottityyppi						Pottien tulo
	0	1	2	3	4	5	
1	1,1	2,2	3,5	5,0	7,0	10,0	2 965
2	1,1	2,2	3,5	5,0	7,0	11,0	3 261
3	1,1	2,2	3,5	5,0	8,0	11,0	3 727
4	1,1	2,2	3,5	6,0	9,0	11,0	5 031
5	1,1	2,2	3,5	7,0	9,0	11,0	5 870
6	1,1	2,2	4,0	6,0	8,0	11,0	5 111
7	1,1	2,2	4,0	7,0	9,0	11,0	6 708
8	1,1	3,0	4,5	6,0	8,0	11,0	7 841
9	1,1	3,0	5,0	7,0	9,0	11,0	11 435
10	1,1	3,0	5,0	7,0	10,0	11,0	12 705
11	1,1	3,0	5,0	8,0	9,0	11,0	13 068
12	1,1	3,0	5,0	8,0	10,0	11,0	14 520
13	1,5	2,5	4,0	6,0	8,0	11,0	7 920
14	1,5	3,0	4,5	6,5	8,5	11,0	12 307
15	1,5	3,0	5,0	7,0	9,0	11,0	15 593
16	1,5	3,0	5,0	7,0	10,0	11,0	17 325
17	2,0	3,0	4,0	6,0	8,0	11,0	12 672
18	2,0	3,0	5,0	7,0	9,0	11,0	20 790
19	2,0	4,0	6,0	8,0	10,0	11,0	42 240
20	3,0	4,0	6,0	8,0	10,0	11,0	63 360

Taulukko 14: Loton simuloinneissa käytetyt pottitarjoamat

Värikoodauksen perusteella voidaan tehdä merkittävä havainto siitä, että sekä tavallisella että logistisella regressiolla saadut tulokset asettuvat hyvin pitkälle samaan paremmuusjärjestykseen tarkasteltavasta kategoriasta riippumatta. Tulosta voidaan pitää rohkaisevana simulointitulosten luotettavuuden kannalta.

Tarkasteltaessa edellä kuvatun pisteytysfunktion avulla arvoitettuja pottitarjoamia, voidaan havaita, että neljä viidestä parhaiten pisteytetystä pottitarjoamasta ovat samat regressiomallien simulointitulosten osalta. Tulos tukee entisestään sitä, että simulointimallien avulla on mahdollista vertailla eri pottitarjoamamallien toimivuutta Loton pottienhallinnan osalta.

Tarkasteltaessa niin myyntien keskiarvoa kuin mediaaniakin voidaan tehdä suhteellisen odotettu ja tämän työn tulosten kannalta merkittävä havainto siitä, että keskiarvoltaan ja mediaaniltaan suurimman myynnin omaavat pottitarjoamat ovat niitä tarjoamia, joissa Loton potti kasvaa mahdollisimman aggressiivisesti. Eräänä poikkeuksena tähän voidaan huomata olevan pottitarjoama 11, joka ei ole pottien tulon mukaan sieltä aggressiivisimmasta päästä, mutta omaa kuitenkin yhden korkeimmista vuositasen keskiarvoisista myynneistä. Keskimyyntien osalta voidaan havaita, että pottitarjoamien välillä on hyvinkin merkittäviä eroja vuositasolla. Suurimpien ja alhai-

simpien keskimyyntien erot pottitarjoamien välillä ovat noin 90 miljoonan euron suuruusluokkaa. Eroa voidaan pitää odotettuna, sillä aggressiivisempien pottitarjoamien voidaan olettaa generoivan myyntiä erittäin tehokkaasti verrattuna maltillisempiin tarjoamiin. Todellisuudessa pottitarjoamien välisen myynnin erot luultavasti olisivat kuitenkin alhaisempia kuin 90 miljoonaa euroa. Tämä johtuu siitä, että hyvin aggressiiviset pottitarjoamat todennäköisesti alkaisivat ajan myötä tavanomaista voimakkaammin inflatoitumaan, joten niiden vuositason keskimyyntit olisivat alhaisempia kuin mitä simuloinneissa on tulokseksi tullut. Vaikka inflatoituminen aggressiivisimpien tarjoamien myyntiin varmasti todellisuudessa vaikuttaisi, niin vuositasolla eri pottitarjoamien välisten keskimyyntien ero on varmasti kymmenien miljoonien eurojen suuruusluokkaa. Taulukosta 13 voidaan havaita, että tavallisella regressiolla saadut myynnin simulointitulokset ovat järjestään hieman suurempia kuin logistisella regressiolla saadut tulokset. Tulos on sinällään odotettu, sillä kuten edellä regressiodiagnostiikan osalta huomattiin, on tavallisen regressiomallien myynnin estimaatit varsinkin suuremmilla poteilla pääsääntöisesti korkeammat kuin logistisella regressiolla. Tämä ei kuitenkaan olennaisesti vaikuta simulointitulosten tulkintaan, sillä tässä yhteydessä olennaisempaa on vertailla eri pottitarjoamien suhtautumista toisiinsa, ei niinkään keskimyyntien absoluuttisia arvoja.

Tarkasteltaessa myyntien variansseja voidaan huomata, että ne ovat suurempia sellaisilla pottitarjoamilla, joissa jossain välissä tehdään suurempi hyppäys pottien suuruuksien välillä siirryttäessä pottityypistä toiseen. Tulosta voidaan pitää odotettuna, sillä mitä pienempiä hyppäyksiä pottien suuruuksissa tehdään, sitä vähemmän vuositasolla sattuma vaikuttaa pottien suuruuksiin ja edelleen vuositason myyntiin.

Eräs tämän työn mielenkiintoisimmista tuloksista on havaittavissa tarkasteltaessa myyntien persentilejä. Tarkasteltaessa alhaisimman 5 ja 95 prosentin persentilejä huomataan, että pottitarjoamien sisällä vuositason myynnin vaihtelut ovat 60 - 90 miljoonan euron suuruusluokkaa. Vaikka tässäkin tuloksessa pottien inflatoituminen jonkin verran todennäköisesti edellä mainittuun suuruusluokkaan vaikuttaa, puhutaan varmasti jokaisen pottitarjoaman osalta kymmenien miljoonien erosta parhaimpien ja huonoimpien pottituurien välillä (5 ja 95 prosentin persentilit). Tarkasteltaessa 25 ja 75 prosentin persentilejä huomataan, että tässäkin tapauksessa simulointien perusteella vuositason myyntien erot ovat eri pottitarjoamilla 35 miljoonan euron suuruusluokkaa. Pottien inflatoituminen tähänkin todennäköisesti hieman vaikuttaa, mutta uskoisin 25 ja 75 prosentin persentiliiden eron olevaan vähintään 20 miljoonan euron suuruusluokkaa. Edellä kuvattuja vuositason vaihteluita voidaan pitää erittäin mielenkiintoisena tuloksena koskien Loton vuositason myyntien käyttäytymistä. Jos vuositasolla sattuu esiintymään poikkeuksellisen paljon tai vähän suuria potteja, vaihtelee Loton vuositason myynti erittäin merkittävästi. Vaihtelu on huomattavasti suurempaa, mitä olisin etukäteen odottanut. Persentiliiden osalta voidaan havaita, että tavallisella regressiomallilla persentiliiden absoluuttiset arvot ovat odotetusti korkeammat kuin logistisella regressiolla. Olennaista on kuitenkin huomata se,

että kummallakin mallilla persentiilien erot ovat samaa suuruusluokkaa. Uskoisin siis, että edellä kuvaamani vuositason vaihtelut kuvastavat pottien inflatoitumista lukuun ottamatta hyvin sitä suuruusluokkaa, kuinka Loton myynti vuositasolla vaihtelee pottituurista riippuen.

Taulukossa 13 on myös esitetty vuositasolla eri pottitarjoamien voittorahaston absoluuttisen kulutuksen keskiarvo, mediaani ja varianssi. Tässä taulukossa ei ole huomioitu sitä, että kukin pottitarjoama Loton sääntöjen mukaan myös kasvattaa voittorahastoa (tässä yhteydessä 0,9% liikevaihdosta). Tarkasteltaessa eri pottitarjoamien vuositason voittorahaston kulutuksia voidaan havaita, että niiden välillä on hyvinkin merkittäviä eroja. Alhaisimmilla vuositason käytön keskiarvo on noin miljoonan euron suuruusluokkaa, kun taas aggressiivisimmilla malleilla vuositason käyttö on jopa 40 miljoonan euron tasolla. Tarkasteltaessa niin tavallisen kuin logistisen regression simulointien voittorahaston käyttöä, voidaan selkeästi havaita että logistisen regression simuloinneissa voittorahaston käyttö on aina suurempaa kuin tavallisen regression simuloinneissa. Tulos on odotettu, sillä kuten jo edellä todettiin, logistisen regression myyntien estimaatit ovat etenkin suuremmilla poteilla alhaisempia kuin tavallisella regressiolla. Myyntiestimaattien ero vaikuttaa suoraan siihen, kuinka paljon mil-läkin kierroksella voittorahastoa potin kattamiseen tarvitaan.

Tässä yhteydessä on syytä huomata, että erittäin suuret voittorahaston kulutuksen odotusarvot tarkoittavat sitä, ettei kyseistä pottitarjousmallia voitaisi käytännössä kokonaista vuotta Lotossa hyödyntää. Vuositasolla voidaan karkeasti sanoa, ettei Lotossa voida voittorahastoa käyttää kuin maksimissaan 20 miljoonan euron edestä, eikä tätäkään tasoa voitaisi ylläpitää vuodesta toiseen. Näin ollen ainakin pottitarjoamat 19 ja 20 eivät olisi sellaisia, joita voitaisiin yhtäjaksoisesti vuoden ajan hyödyntää (ellei esimerkiksi Loton päävoittoluokan osuutta voitonjaosta kasvatettaisi). Tätä tulosta voidaan myös pitää simulointien osalta hyvinkin merkittävänä.

Tarkasteltaessa voittorahaston kulutuksen variansseja voidaan huomata, että varianssi selvästi kasvaa voittorahaston kulutuksen kasvun mukana. Tulosta voidaan pitää odotettuna, sillä kaikkein aggressiivisimmillakin malleilla kulutus voi sattuman myötä olla hyvin alhaista.

Kuvassa 14 on esitetty edellä kuvatun pisteytysfunktion (19) mukaan viidelle parhaalle tavallisen regression pottitarjousmallille vuotuisen myynnin jakauma. Kuvassa on myös esitetty vastaaville malleille logistisen regression avulla simuloidut myyntien jakaumat. Tavallisen regression tulokset ovat kuvan vasemmalla puolella ja logistisen regression tuloksen kuvan oikealla puolella. Tässä kohtaa ei katsottu tarpeelliseksi määrittää vastaavaa kuvaa parhaille logistisen regression avulla simuloiduille myynneille, sillä lopputuloksena olisivat olleet lähes identtiset kuvat pottitarjoamien osalta.

Kuvasta 14 voidaan havaita se, kuinka tavallisen regression parhaiden pottitarjousten myyntien jakaumat eroavat hyvin merkittävästi toisistaan. Vastaavasti voidaan sama huomata myös logististen regressiomallien avulla tehdyistä simulointituloksista. Esimerkiksi mallin 1-1 alimmat vuotuiset myyntitasot alkavat noin 420 miljoonan kohdalta, kun taas mallilla 1-17 alhaisimmat vuotuiset myynnit ovat 450 miljoonan suuruusluokkaa. Tässä voidaan todeta, että pisteytysfunktio on kuitenkin nostanut mallin 1-1 paremmaksi kuin mallin 1-17 johtuen esimerkiksi mallin 1-17 suuremmasta odotusarvoisesta voittorahastonkäytöstä suhteessa voittorahaston kartuttamiseen.

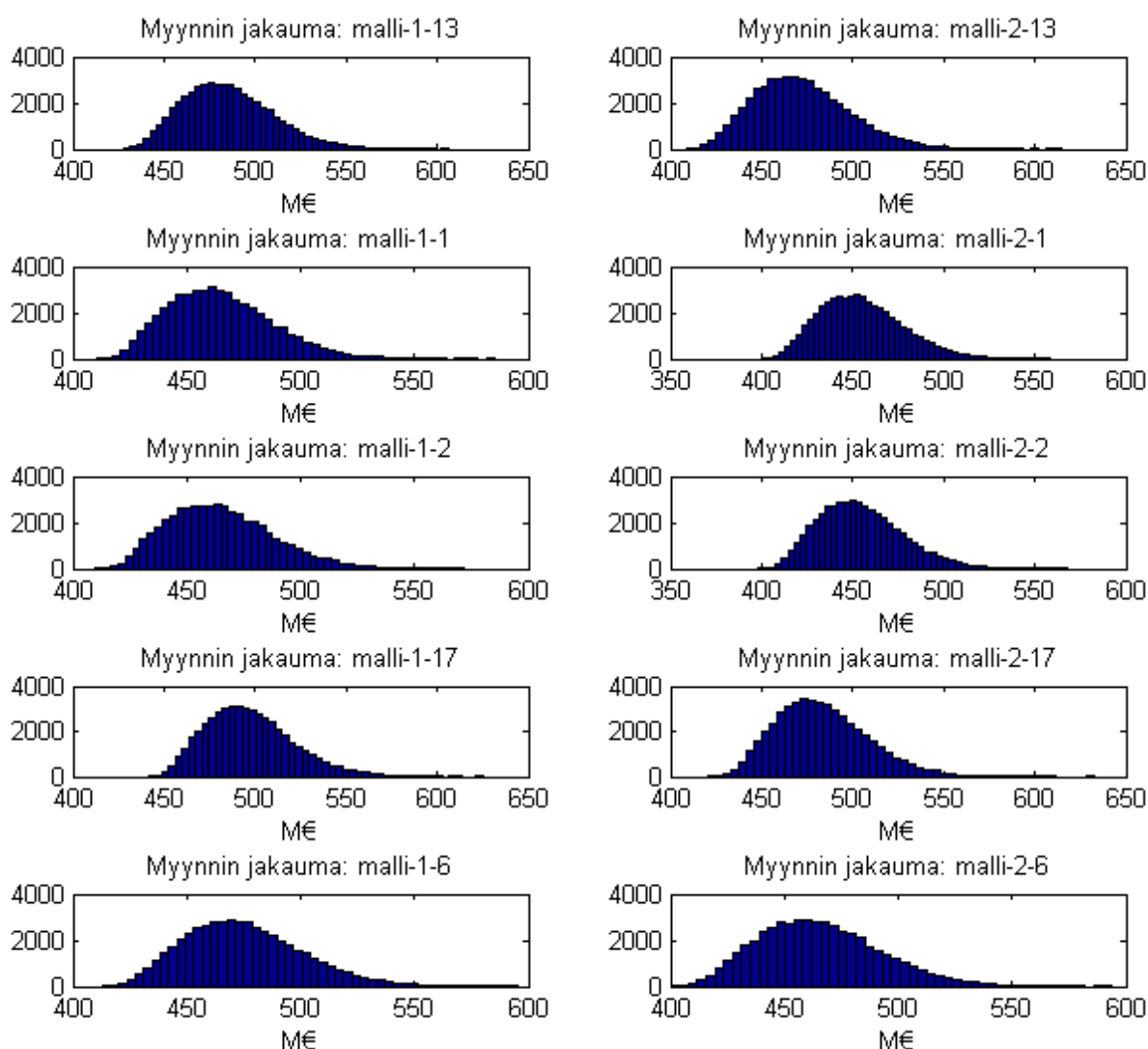
Mielenkiintoista on myös verrata, kuinka tavallisella regressiomallilla laaditut simuloinnit poikkeavat logistisen regressiomallin simulointituloksista. Selkeimpänä eroavaisuutena voidaan nähdä, että logistisella regressiomallilla on kautta linjan saatu hieman alhaisempia vuotuisia myyntejä. Tämä voidaan havaita vertailemalla esimerkiksi alhaisimpia vuotuisia myyntituloksia sekä tarkastelemalla simulointitulosten huippukohtia. Tulos on odotettu johtuen jo edellä mainitusta havainnosta, että logistisen regression mallit estimoivat varsinkin suurilla poteilla kertapeliä myynnin alhaisemmaksi kuin tavallisen regression mallit.

Vertailtaessa simulointitulosten jakaumien muotoja tavallisen ja logistisen regression välillä ei voida mielestäni nähdä merkittäviä eroja, mitä voidaan tämän työn tulosten luotettavuuden kannalta pitää rohkaisevana asiana. Vertailtaessa jakaumien muotoja niin tavallisen kuin logistisen regression sisällä voidaan eroja jo kuitenkin havaita. Vaikuttaisi siltä, että tietyillä pottitarjousmalleilla simulointitulosten kuvio on kaapeampi kuin toisilla. Tämä voidaan havaita esimerkiksi mallien 1-1 ja 1-6 sekä mallien 2-1 ja 2-6 simulointitulosten kuvien välillä. Tulos on sinällään odotettu, sillä taulukon 13 mukaan edellä kuvatuilla malleilla on selkeä ero simulointitulosten varianssissa.

Simulointitulosten jakaumista voidaan vielä havaita se mielenkiintoinen tulos, että vaikuttaisi siltä, että jakaumat ovat positiivisesti vinoja vuotuisen myynnin osalta kautta linjan. Tämän vahvistaa taulukon 13 tarkastelu, josta nähdään, että jakaumien mediaanit ovat järjestään keskiarvoja alhaisempia. Tämä voidaan tulkita siten, että loton vuotuinen myynti jää yli 50 % kerroista hieman simulointimallien keskimääräistä myyntiä alhaisemmaksi. Tätä voidaan mielestäni pitää mielenkiintoisena havaintona.

Merkittävimpiä havaintona myyntien jakaumista voidaan tehdä se, että pienempien ja suurimpien vuotuisen myyntien välinen ero on todella merkittävä. Kaikilla pottitarjousmalleilla puhutaan yli 100 miljoonan euron erosta. Silmämääräisesti tarkasteltuna voidaan havaita, että 50 miljoonan vuotuisen myynnin ero ei ole edes mitenkään hirveä epätodennäköinen. Tulosta voidaan pitää erittäin merkittävänä ja se kuvaa erinomaisesti lottopelien luonnetta. Vastaavia tuloksia saatiin myös myyntien persentilejä tarkastelemalla. Mikäli vuodenajanjaksolle sattuu osumaan keskimääräistä enemmän suuria potteja, kasvaa vuotuisen myynnin suuruus todella voimakkaasti.

Toisaalta jos potit pysyvät vuoden ajanjaksolla suhteellisen alhaisina voi myynti jäädä merkittävästikin odotusarvoisesta tasosta. Tämä tulos kuvaa mielestäni hyvin Loton pelaamisen luonnetta väestön keskuudessa. Kun suuria potteja on pelaajille tarjolla, innostuu yhä useampi pelaaja pelaamaan. Toisaalta mikäli potit jäävät keskimääräistä alhaisemmiksi, pelaavat peliä lottopeleille uskollisimmat asiakkaat alhaisista poteista huolimatta.



Kuva 14: Vuotuisen myynnin jakauma pisteytysfunktion mukaan viidelle parhaalle tavallisen regression mallille

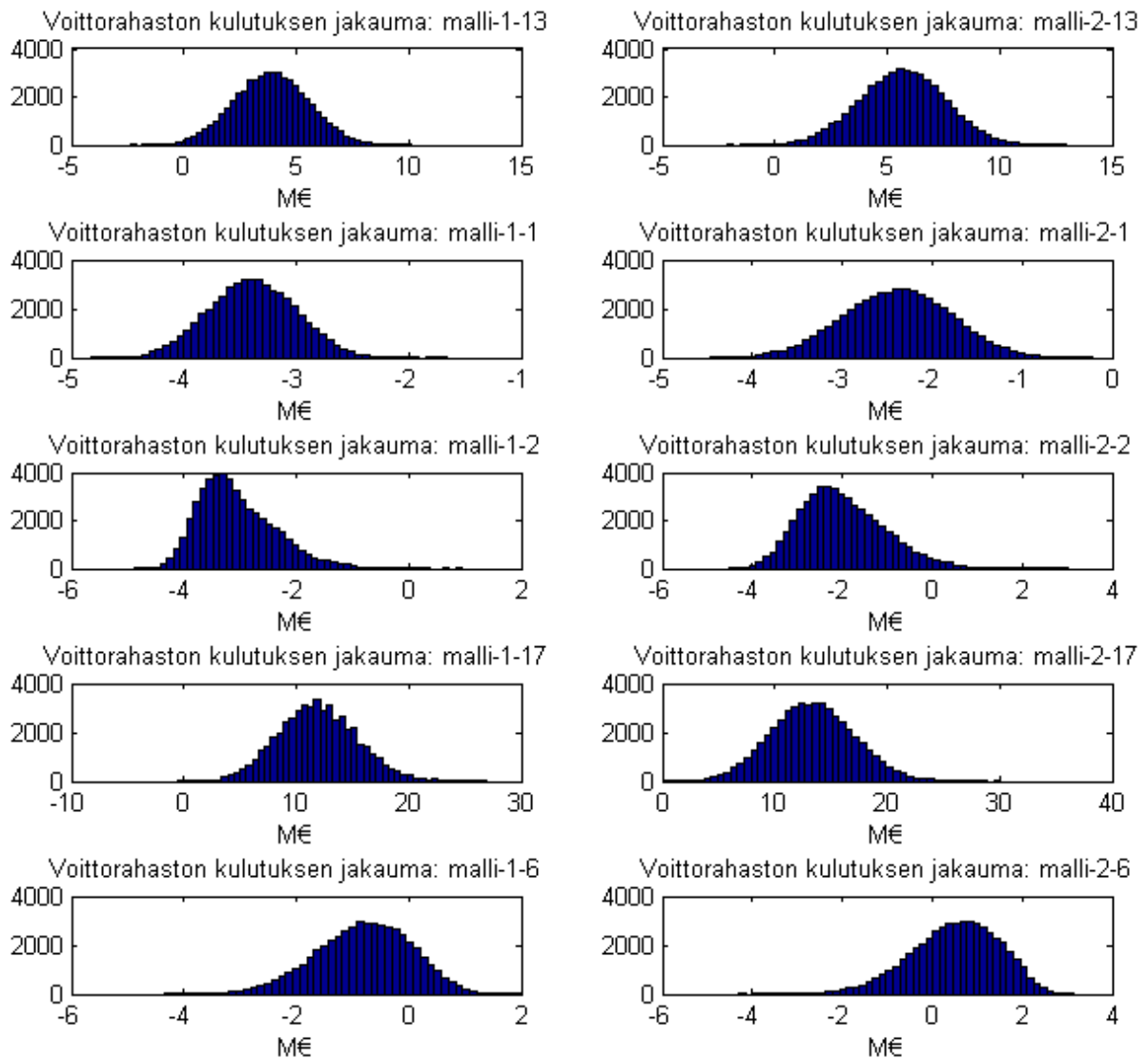
Kuvassa 15 on kuvattu voittorahaston käytön kulutusjakaumia vastaaville pottitarjouksmalleille, mitä edellä käsiteltiin kuvan 14 osalta.

Kuvasta 15 voidaan havaita se, kuinka merkittävästi voittorahaston kulutuksen jakaumat vaihtelevat eri mallien välillä niin tavallisen kuin logistisen regressionkin välillä. Osalla malleista kulutusjakauma on selkeästi positiivisen puolella kautta linjan, kun taas osalla malleista jakauma jää kauttaaltaan negatiiviseksi. Negatiivinen jakauma tulkitaan tässä yhteydessä siten, että käytettävä pottitarjouksmalli kuluttaa vuotui-

sesti vähemmän voittorahastoa mitä se Loton sääntöjen mukaan sitä kartuttaa (tämän työn osalta 0,009€ kutakin pelattua euroa kohden). Tässä yhteydessä on syytä huomauttaa, että taulukossa 13 ei huomioitu pottitarjousmallien voittorahaston kartuttamista, mistä johtuen esimerkiksi taulukon 13 keskiarvoinen voittorahaston kulutus eroaa selkeästi kuvan 15 tuloksista.

Vertailtaessa tavallisella ja logistisella regressiomallilla saatuja jakaumia voidaan niiden havaita olevan hyvin samankaltaisia suuruusluokkiensa osalta. Tätä tulosta voidaan pitää mielestäni hyvänä simulointitulosten luotettavuuden osalta. Suurimman kulutuksen omaava malli on selvästi malli 17 ja pienimmät kulutukset malleilla 1, 2, ja 6.

Vertailtaessa tavallisen ja logistisen regression avulla simuloitujen jakaumien muotoja voidaan havaita, että ne ovat pääsääntöisesti hyvin samankaltaisia. Suurimpana poikkeuksena ehkä malli 2, jossa tavanomaisella regressiolla saatu jakauma vaikuttaisi olevan huipukkaampi kuin logistisella regressiolla saatu jakauma. Tätä ei kuitenkaan voida vahvistaa taulukon 13 avulla, sillä simuloinneissa en määrittänyt jakaumien huipukkuutta.



Kuva 15: Voittorahaston kulutuksen jakauma pisteytysfunktion mukaan viidelle parhaalle tavallisen regression mallille

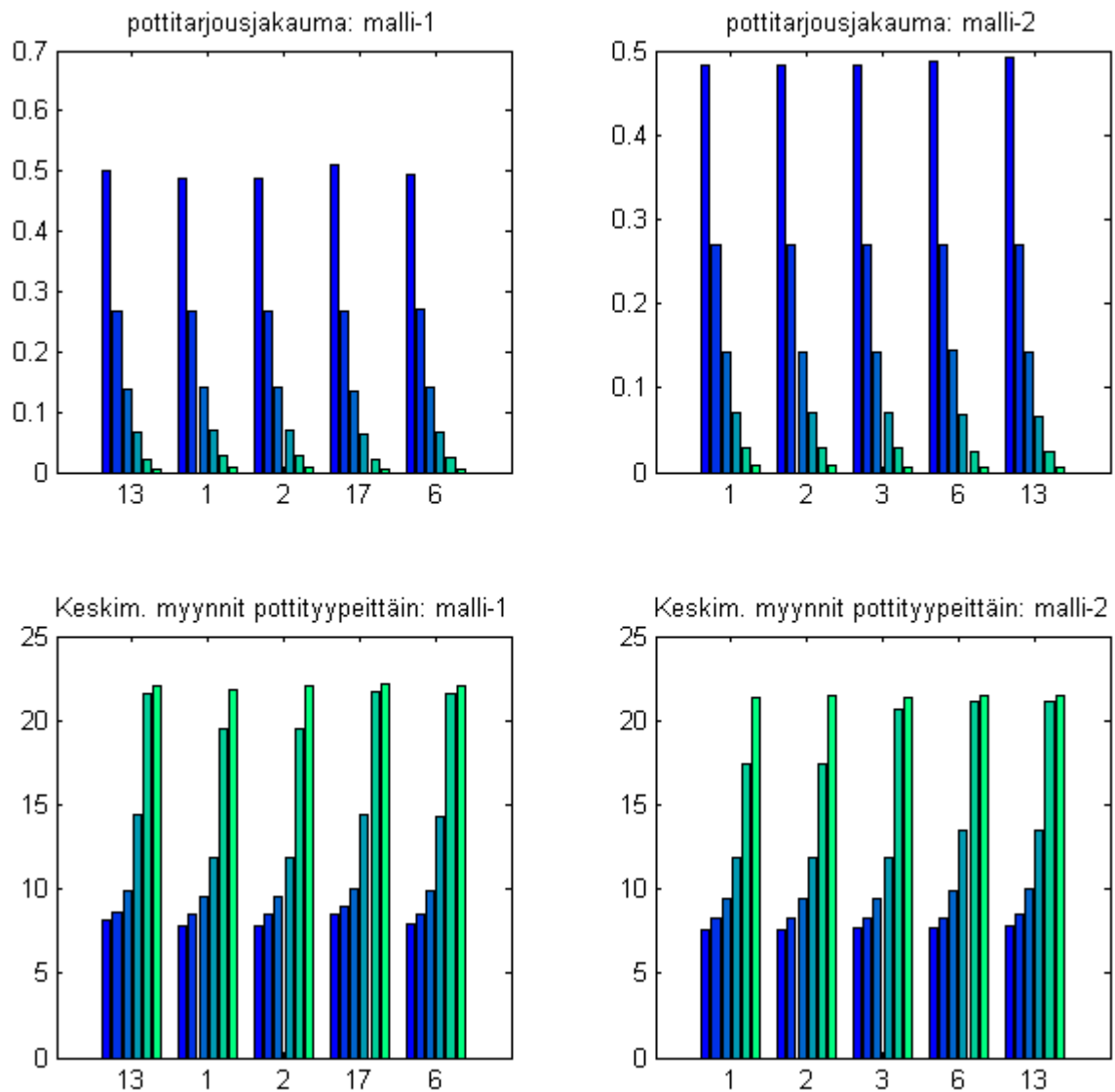
Kuvissa 14 ja 15 esitetyistä simulointitulosten jakaumista voidaan tehdä seuraava merkittävä huomio. Negatiivinen jakauma voittorahastonkäytöllä ei välttämättä ole pottitarjousmallin kannalta huono asia. Kuten tässä työssä on edellä kuvattu, siirrettään simuloinneissa jokaisesta Lottoon pelatusta eurosta 0,009€ voittorahastoon ja nämä rahat on käytettävä nimenomaan Loton voittoina (41/2011 alkaen 0,020€ voittorahastoon peliriviä kohden). Näin ollen, mikäli voittorahaston käyttö olisi oletusarvoisesti negatiivista, kerryttäisi pottitarjousjakauma enemmän voittorahastovaroja mitä se kuluttaisi vuoden aikana. Tällaisessa tilanteessa Lotossa tarjottaisiin aika ajoin tehdä korkeampia potteja kuin pottitarjousmallissa on kuvattu. Tämä oletusarvoisesti kasvattaisi myyntiä, joten tosiasiaa mallien 1, 2 ja 6 myynninjakauma siirtyisi jonkin verran oikealle. Tällaista ominaisuutta ei kuitenkaan tähän simulaattoriin rakennettu, sillä sen toteuttaminen olisi hyvin haastavaa. Tässä yhteydessä voidaan vaan todeta, että negatiivinen voittorahaston oletusarvoinen käyttö ei käytännössä

olisi pottitarjousmallin osalta huono asia, vaan sitä voitaisiin jopa pitää pottitarjousmallin vahvuutena.

Kuvassa 16 on esitetty pottitarjousjakauma (kuvan yläosa) sekä keskimääräiset myynnit pottityypeittäin viidelle parhaalle pottitarjousmallille niin tavallisen kuin logistisen regression avulla simuloitujen tulosten osalta.

Kuvasta 16 voidaan vuotuisten pottitarjousjakaumien osalta tehdä seuraavia havaintoja (pottitarjousjakaumalla tarkoitetaan tässä periaatteessa pottityyppijakaumaa). Noin puolet kierroksista on ollut niin sanottuja peruskierroksia eli lottopotteja, joita edeltävällä kierroksella on löytynyt Lotossa päävoitto. Tavallisen regression osalta voidaan tehdä mielenkiintoinen havainto siitä, että mallilla 17 näyttäisi peruskierrosten osuus olevan kohtalaisen selvästi suurempi kuin muilla pottitarjousmalleilla. Tämä selittynee sillä, että pottitarjousmalli 17 omaa selkeästi nopeammin kasvavat pottit kuin muut tarkastelun pottitarjousmallit (ks. taulukko 14). Näin ollen muilla kuin peruskierroksilla pelaaminen lisääntyy nopeammin, mistä seuraa että todennäköisyys että päävoittajia löytyy kasvaa. Millään logistisen regression parhaista pottitarjousjakaumista ei ole näin selvästi vastaavaa ilmiötä havaittavissa. Muuten pottityyppijakaumat näyttäisivät olevan suhteellisen identtiset pieniä vaihteluita lukuun ottamatta. Tämä voidaan tulkita siten, että vaikka pottitarjousjakaumat keskenään vaihtelevat, niin pelaamisen muutos (ja näin ollen Loton riviavaruuden täyttöasteen muutos) eri pottityyppien välillä on kuitenkin suhteellisesti niin vähäistä, ettei se merkittävästi vaikuta vuositasolla eri pottityyppien suhteellisen osuuteen. Tätä voidaan pitää hyvin merkittävänä Loton dynamiikkaa kuvaavana tuloksena.

Keskimääräisten myyntien osalta voidaan havaita, että pottityyppien 0-2 osalta eri pottitarjousmallien väliset erot ovat hyvin vähäisiä (kolme ensimmäistä pylvästä) niin tavallisella kuin logistisella regressiolla muodostetuilla malleilla. Pottityyppien 3-5 osalta pottitarjoamien välillä on jo merkittävämpiä eroja keskimääräisten myyntien osalta. Kuvaajien avulla voidaan identifioida, mistä selittäjistä tiettyjen pottitarjoamien korkeampi vuotuinen odotusarvo muodostuu. Kuten edellä todettiin, ei eri pottityyppien vuotuisten esiintymisjakaumien välillä ole kovinkaan merkittäviä eroja eri pottitarjoamien välillä. Näin ollen vuotuisten myyntien odotusarvojen erot selittyvät valtaosin keskimääräisten myyntien välisillä eroilla. Esimerkiksi pottitarjouksilla 6, 13 ja 17 keskimääräiset myynnit ovat pottityypeillä 3-4 selkeästi suurempia kuin esimerkiksi pottitarjouksilla 1 ja 2. Tämä selittyy suuremmilla poteilla pottityyppien 3 ja 4 osalta. Tätä seikkaa voidaan myös pitää eräänä simulointitulosten merkittävimpänä havaintona.



Kuva 16: Pottitarjousjakaumat ja keskimääräiset myynnit pottityypeittäin viidelle parhaalle pottitarjousmallille

8.3 LOTON POTTIEN VAIKUTUS MYyntIIN JA SEN MUUTTUMINEN AJASSA

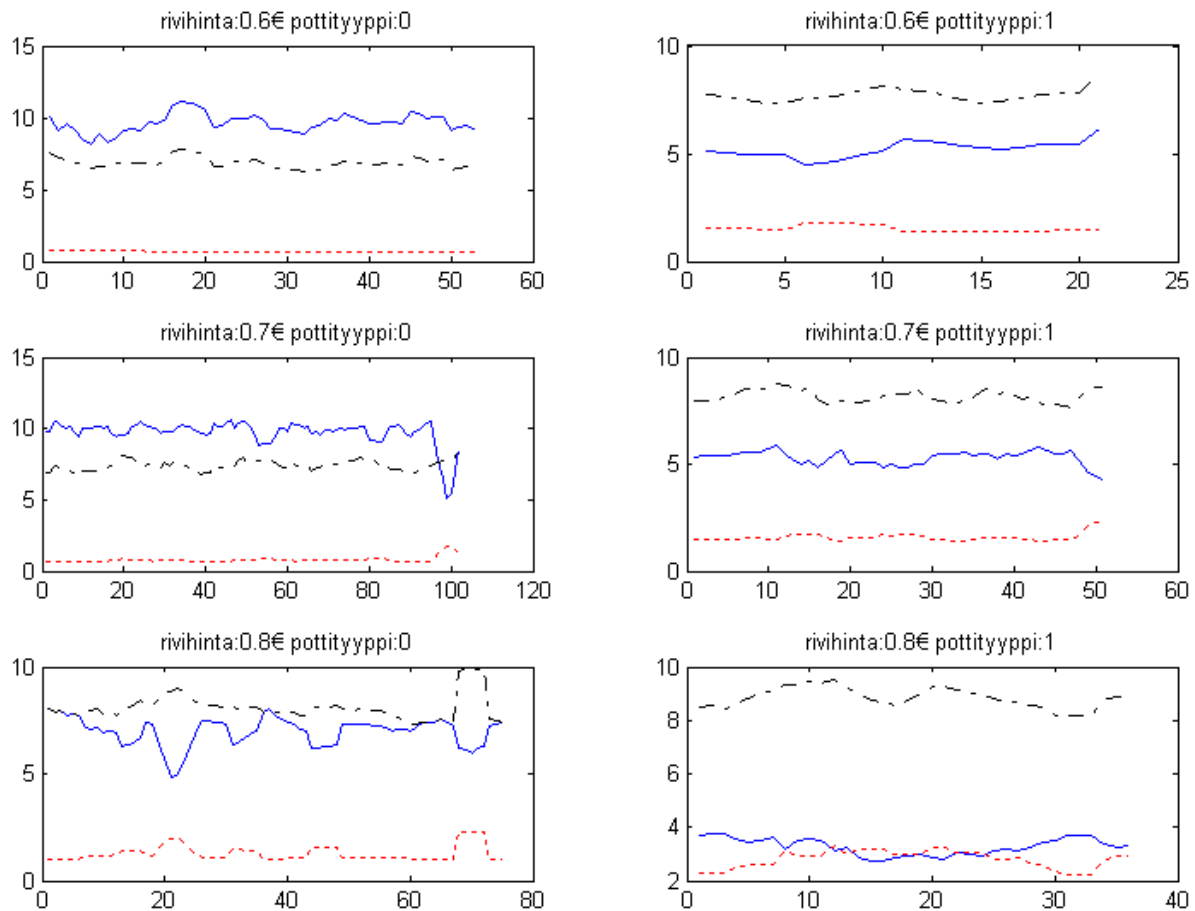
Kuvissa 17 ja 18 on esitetty graafisesti pottien vaikutusta Loton myyntiin ja sen muuttumista ajassa. Datana tässä käytettiin Loton myynti-, pottityyppi- ja pottitieto- ja ajalta 1/2002 - 52/2010. Pottien vaikutusta myyntiin tarkasteltiin määrittämällä erikseen kullekin pottityypille 0-3 niin sanottu myyntisuhde. Myyntisuhdeella tarkoitetaan tässä yhteydessä Loton myynnin suhdetta potin suuruuteen. Tästä myyntisuhdesta muodostettiin viiden viikon liukuva keskiarvo pottityypeille 0 ja 1, sekä kahden viikon liukuva keskiarvo pottityypeille 2 ja 3. Kuvissa 17 ja 18 edellä kuvattu myyntisuhde on esitetty sinisen viivan avulla. Myyntisuhteen lisäksi kuvissa on kuvattu vastaavilla liukuvilla keskiarvoilla niin Loton euromääräinen myynti (musta viiva) kuin myös Loton potti (punainen viiva). Edellä kuvatut myyntisuhteet, myynnit ja Loton potti on kuvissa kuvattu erikseen vielä hinnankorotusten osalta, jotta hinnan-

korotuksen vaikutusta pystytään myyntisuhteeseen, myyntiin ja pottien suuruuteen arvioimaan. Pottityyppien 4 ja 5 osalta vastaavia tarkasteluja ei suoritettu vähäisen datamaarän johdosta.

Kuvassa 17 on esitetty myyntisuhteen kehitys pottityyppien 0 ja 1 osalta rivihinnoitain. Myyntisuhdetta tarkasteltaessa peruskierrosten (pottityyppi 0) sekä rivihintojen 0,60€ ja 0,70€ osalta myyntisuhde on ollut noin 10, eikä myyntisuhteen trendeissä ole havaittavissa selkeää muutosta. Yhdessä suhteellisen stabiilin myyntikäyrän ohella (noin 7,5M€) tämä voidaan tulkita siten että Loton pelaaminen peruskierroksilla oli edellä mainituilla rivi hinnoilla suhteellisen stabiilia, eikä merkittäviä muutoksia pelaamisen trendeissä ollut havaittavissa. Rivi hinnan 0,80€ osalta voidaan havaita, että myyntisuhteen absoluuttinen arvo on tippunut noin kahdeksaan. Tämä selittyy pääosin sillä, että rivi hinnan noustessa 0,70 eurosta 0,80 euroon päävoittoluokan osuutta kasvatettiin 10 prosentista 13,6 prosenttiin. Päävoittoluokan prosenttiosuuden muutos kasvatti Loton potteja edellä mainittujen osuuksien suhteessa. Vaikka myyntisuhteen absoluuttinen arvo laski, voidaan absoluuttisen myynnin kehityksestä havaita, että peruskierrosten myynti nousi selvästi noin 8 miljoonaan euroon. Mielenkiintoisena havaintona voidaan pitää sitä, että kun Loton potteja on nostettu (punainen viiva) on euromääräinen myynti selvästi kasvanut, mutta myyntisuhde puolestaan tippunut, kuten on odotettavissa. Korkeampien pottien vaikutuksista voidaan havaita, että peruskierrosten osalta itse asiassa myyntisuhde on jopa hieman parantunut ajan myötä rivi hinnan 0,80€ osalta (havainnot 43-46 & 68-72 vs. havainnot 20-24). Tulosta voidaan pitää mielenkiintoisena ja se voidaan tulkita niin, että ainakaan tuolloin ei merkittävä inflatoitumista Loton peruskierrosten poteissa tapahtunut.

Pottityypin 1 tuloksista voidaan havaita samankaltaisia asioita kuin peruskierrosten osalta. Myyntisuhteet 0,60€:n ja 0,70€:n rivi hintojen osalta ovat hyvin lähellä toisiaan, absoluuttisen arvon ollessa viiden yläpuolella. Myyntisuhteen voidaan havaita olevan hyvinkin stabiili, 0,60€:n rivi hinnan tapauksessa jopa hieman kasvava. Näin ollen pottien inflatoitumista ei näissäkään tapauksissa ole havaittavissa. 0,80€:n rivi hinnalla myyntisuhteen absoluuttisessa arvossa tapahtuu selvä lasku. Laskun voidaan olettaa johtuvan pääasiassa samasta syystä kuin peruskierrosten tapauksessa, eli päävoittoluokan osuuden kasvattamisesta. 0,80€:n rivi hinnan osalta voidaan havaita että pottien vaihtelu on ollut hyvin suurta. Tarkasteltaessa myyntisuhdetta ja pottien vaihtelua, voidaan kuitenkin huomata, ettei mitään selkeää laskua myyntisuhteessa ole ajassa havaittavissa. Näin ollen voidaan todeta, ettei pottien inflatoitumista ole havaittavissa. Absoluuttisten myyntien osalta voidaan nähdä, että myynnit ovat korkeimmalla tasolla 0,80€:n rivi hinnan tapauksessa. Myynnit myös selkeästi vaihtelevat pottien suuruuden mukaan, mitä voidaan pitää odotettavissa olevana tuloksena.

Kuvan 17 perusteella voidaan sanoa että ainakaan näiden kuvien perusteella ei havaintodatasta ollut löydettävissä pottien inflatoitumista tukevia tekijöitä pottityyp-
pien 0 ja 1 osalta.

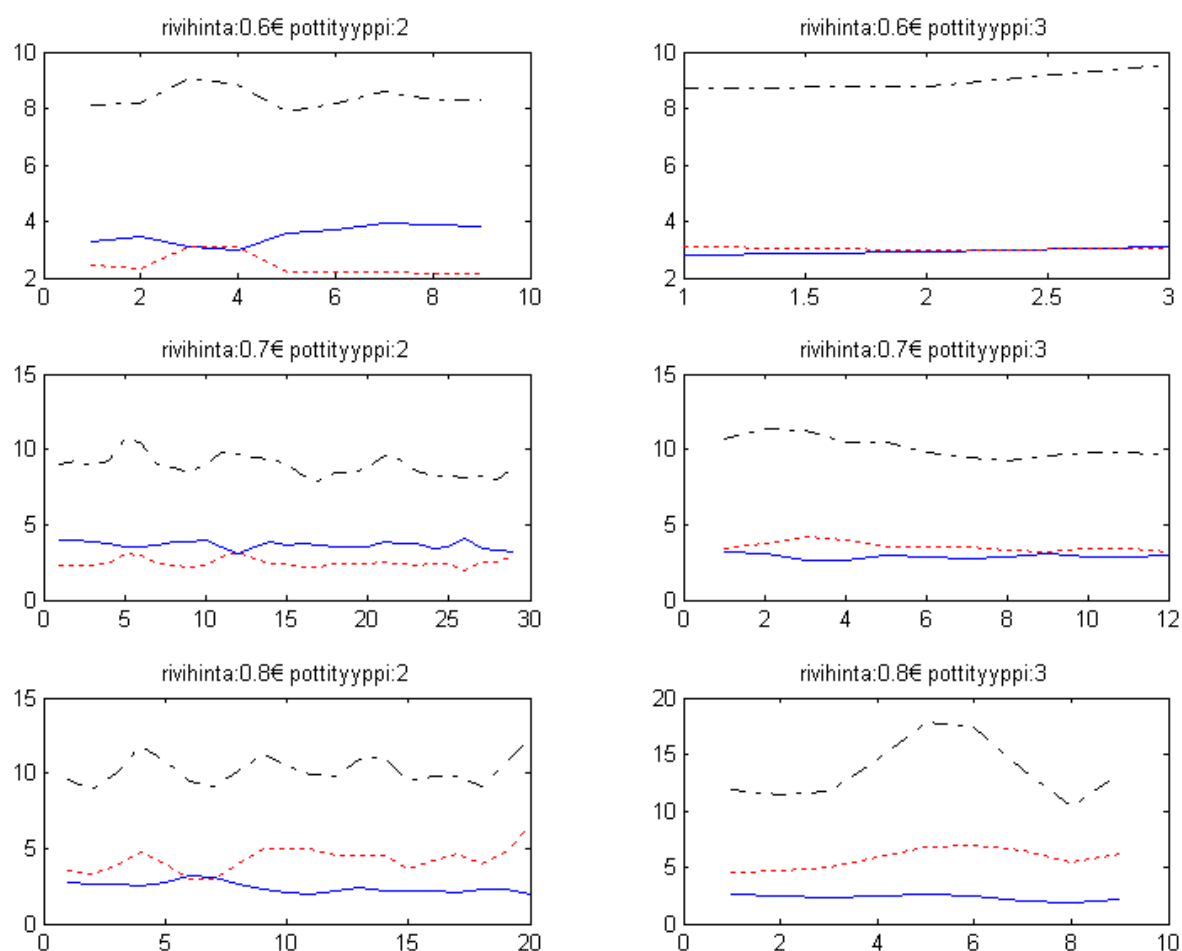


Kuva 17: Myyntisuhteet, myynnit ja potit pottityypeille 0 ja 1

Kuvassa 18 on esitetty rivi hinnoittain myyntisuhteet, potit ja myynnit pottityypeille 2 ja 3. Pottityypin 2 osalta voidaan havaita, ettei myyntisuhteessa ole tapahtunut laskua ajan kuluessa rivi hintojen 0,60€ ja 0,70€ osalta (rivi hinnan 0,60€ osalta jopa selvää nousua). Rivi hinnan 0,80€ osalta voidaan havaita, että myyntisuhteessa saattaisi olla havaittavissa pientä laskua havaintojen 8-17 välillä. Tätä laskua ei voida selittää pottivaihteluilla. Toisaalta lasku on hyvin maltillista, joten mistään merkittävästä inflaatiosta ei voida puhua. Absoluuttisissa myyntimäärissä rivi hinnalla 0,80€ on saavutettu selkeästi suurimmat myynnit. Vaikuttaisi yleisesti ottaen siltä, että pottityypin kaksi osalta rivi hintojen korotus on yleisesti kasvattanut myyntien tasoa (toki tähän vaikuttaa myös varmasti kasvaneet potit).

Pottityypin 3 tulosten osalta on jo hankalampi tehdä mitään johtopäätöksiä suhteellisen alhaisen havaintomäärän takia. Kuvasta 18 voidaan kuitenkin havaita, että myyntisuhteessa ei ole havaittavissa mitään selkeää inflatoitumista rivi hintojen 0,60

€ ja 0,70€ osalta. Rivihinnan 0,80€ osalta voidaan ehkä pienekköä inflatoitumista havaita, mutta tulkintaa tässä tapauksessa vaikeuttaa selkeästi vaihdelleet potit. Absoluuttisissa myyntimäärissä voidaan tehdä vastaavia havaintoja kuin pottityypin 2 tapauksessa. Pottityypin 3 myynnit ovat kasvaneet selkeästi rivihinnan mukana. 0,80€:n rivihinnan tapauksessa voidaan havaita, että kuvan keskivaiheilla olevilla isoilla poteilla on saavutettu todella korkeita myyntilukuja. Tätä selittää pääosin se, että tuolle tarkastelujaksolle osui useampi Loton niin kutsuttu ennätyspottikierros, jolloin Loton myynti oli erittäin korkealla tasolla.



Kuva 18: Myyntisuhteet, myynnit ja potit pottityypeille 2 & 3

Yleisesti ottaen kuvien 17 ja 18 osalta voidaan todeta, että mitään selkeää pottien inflatoitumista ei tällä aineistolla ja tällä tutkimustavalla ole pottityyppien 0-3 tapauksissa havaittavissa. Alempien pottityyppien osalta tulosta voidaan pitää luonnollisena, sillä niiden tapauksessa pelaaminen on enemmänkin tapapelaamista (tätä on selvitetty esimerkiksi Veikkauksen asiakastutkimuksissa). Tulosta ei voida kuitenkaan pitää todisteena siitä, ettei inflatoitumista Loton suuremmissa poteissa ajan myötä tapahtuisi. Tarkasteluaineisto ei esimerkiksi ota huomioon pottityyppejä kuin kolmeen saakka. Toisaalta regressiomallien ennustuskyky selkeästi kuvasi Loton pot-

tien inflatotumista. Suurempien pottien inflatoitumista ei kuitenkaan tässä työssä tämän tarkemmin selvitetä ja se jääkin tulevaisuuden tutkimustehtäväksi.

8.4 SIMULOINTIMALLIN USKOTTAVUUS TARKASTELUJA SEKÄ KEHITYSMAHDOLLISUUKSIA

Simulaattorin keskeisimpänä tavoitteena oli vertailla eri pottitarjoamien välisiä vuositasen myynti- ja voittorahaston kulutusjakaumia. Uskoisin, että simulaattorin avulla voidaan suhteellisen luotettavasti vertailla eri pottitarjousmalleja niiden suhteellisen suoriutumisen osalta. Tässä yhteydessä tällä tarkoitetaan sitä, että simulointimallien avulla kyetään vertailemaan sitä miten eri pottitarjousmallit suhtautuvat vuositasolla toisiinsa. Uskoisin myös, että tarkasteltaessa eri pottitarjousmallien keskimääräisten myyntien persenttiilejä, voidaan niiden vuositasen vaihteluista saada simulointimallien avulla saada suhteellisen luotettavaa tietoa. Hieman varauksellisemmin tulee suhtautua kuitenkin sekä myynnin että voittorahaston vuotuisiin kulutusjakauksiin absoluuttisten lukuarvojen osalta.

Suurimmaksi simulaatiomallien heikkoudeksi voidaan eittämättä lukea se, että ne eivät kykene mallintamaan Loton suurien pottien inflatoitumista ajassa. Tämä vaikuttaa suoraan myyntijakaumiin siten, että simuloitujen vuositasen myynnit ovat hyvin todennäköisesti jonkin verran todellisuutta vastaavia korkeampia. Osittain tähän problematiikkaan vastaa se, että suurimmaksi pottityypiksi rajattiin simuloinneissa 5, tosin tämäkään ei täysin tarkasteltavaa ongelmaa ratkaise. Toisena heikkoutena voidaan pitää sitä, että osa pottitarjousmalleista kuluttaa voittorahastoa niin vähän, että vuositasolla kyseisiä pottitarjousmalleja hyödynnettäessä aika ajoin Loton pottina tultaisiin tarjoamaan korkeampia potteja mitä pottitarjousmalliin on kiinnitetty. Tällaisen mekanismin toteuttaminen simulaattoriin on kuitenkin suhteellisen haastavaa, eikä sitä tässä työssä laadittu. Ratkaisu voisi olla esimerkiksi sellainen, että osan ajasta sovellettaisiin tiettyä pottitarjousmallia, jossa voittorahaston odotusarvoinen kulutus on suurempaa.

Jatkokehitysten osalta simulaattoriin voisi pyrkiä rakentamaan mekanismia, joka kykenisi paremmin huomioimaan Loton pottien inflatoitumista. Tämä problematiikka koskee eritoten regressiomallien laatimisvaihetta. Toisena kehityskohteena voidaan nähdä juuri edellä lyhyesti kuvatut niin sanotut älykkäät pottitarjoamat. Tällä tarkoitetaan sitä, että pottitarjoamaa ei kiinnitettäisi yksittäisessä simuloinnissa kokonaan, vaan pottitarjoama voisi muuttua esimerkiksi voittorahaston kulutuksen mukaan.

9 JOHTOPÄÄTÖKSET

Lotto on eittämättä eräs suomalaisen rahapelihistorian merkittävimmistä peleistä. Lotto itsessään omaa jo niin merkittävän roolin suomalaisten rahapelaamisessa, että lähtökohdat tämän työn laatimiseen olivat hyvin mielenkiintoiset.

Työtä tehdessä oli antoisaa perehtyä niin lottopelien historiaan maailmalla kuin Suomessa järjestettävän Loton historiaan sekä rahapelitekniisiin ominaisuuksiin. Mielenkiintoista oli myös nähdä, kuinka lottopelejä ja niiden rahapeliominaisuuksia on tieteellisissä julkaisuissa tutkittu. Vaikka en tätä työtä vastaavaa lähtökohtaa yhdestäkään lottopelien myynnin mallintamista tutkineesta tieteellisestä julkaisusta löytänyt, sain Loton mallinnostyöhön paljon ajatuksia tieteellisten julkaisujen pohjalta.

Tämän työn eräs keskeisimmistä tavoitteista oli muodostaa matemaattiset mallit Loton kestojen kierroskohtaiselle myynnille. Matemaattisiksi malleiksi valikoitui lopulta sekä usean selittäjän tavanomainen että logaritmoitu regressiomalli. Myynnin mallintamisen ohessa tärkeänä tavoitteena oli luoda simulaattori, jonka avulla Loton vuositason myyntejä ja voittorahaston kulutuksia kyettäisiin simuloimaan.

Regressiomallien muodostamisessa onnistuttiin työn tekijän mielestä hyvin. Uskon että valtaosa Loton myyntiin vaikuttavista tekijöistä saatiin tavalla tai toisella sisällytettyä regressiomalleihin. Toki oli joitain selittäjäkandidaatteja, joita ei malleihin saatu mukaan, mutta tämäkin asia tiedostettiin ja se tulee ottaa huomioon tulkittaessa Loton myyntiin vaikuttavia tekijöitä.

Loton myynnin mallintamisen osalta onnistuttiin selvästi parhaiten kertapelien osalta. Tulosta voidaan pitää siinä mielessä odotettuna, että kertapelien pelaamisen voidaan katsoa riippuvan eniten Loton pottitasosta sekä esimerkiksi siitä, kuinka monen kierrokseen ei päävoittoa ole löytynyt. 2, 3 ja 5 viikon kestojen mallinnoksissa onnistuttiin myös oikein mukavasti, mutta 10 viikon kestopelaamisen mallinnus jäi yllättävän heikoksi. Yleisesti ottaen voidaan todeta, että mitä pidemmästä kedestä on kysymys, sitä vähemmän sen pelaamiseen ulkoiset tekijät näyttäisivät vaikuttavan ja pelaaminen on enemmänkin tapapelaamista.

Simulaattorin tärkeimpänä tavoitteena oli kyetä vertailemaan erilaisia pottitarjoamia vuotuisen myynnin ja voittorahaston kulutuksen jakaumien osalta. Tässä tavoitteessa mielestäni onnistuttiin hyvin. Simulointitulosten perusteella voitiin tehdä useita merkittäviä havaintoja Loton pelaamisen dynamiikasta. Mielestäni eräs tämän työn mielenkiintoisimmista tuloksista on se, että vuositasolla Loton myynti vaihtelee kymmeniä miljoonia euroja sen mukaan millaisia potteja vuoden ajanjaksolle sattuu osumaan. Toisena erittäin mielenkiintoisena havaintona voidaan pitää sitä, että vuositasolla tarjotut pottitarjoamat eivät juuri vaikuta pottityyppijakaumien odotusarvoon. Tämä tulos tarkoittaa sitä, että vuositasolla eri pottitarjoamien väliset myyntierot selittyvät pääsääntöisesti pottityyppien välisten myyntien eroilla eikä sillä, että toisella tarjoamalla olisi esimerkiksi selvästi enemmän suuren pottityypin potteja.

Työtä tehdessä identifioitiin myös lukuisia kehitysmahdollisuuksia niin regressiomallien kuin simulaattorin osalta. Regressiomalleihin voisi pyrkiä sisällyttämään tietoa muun muassa kunkin kierroksen säätilasta sekä markkinointitoimenpiteistä. Edellä

kuvatun lisäksi esimerkiksi differenssoinnin avulla saatettaisiin regressiomallien selityskykyä pystyä parantamaan. Eräs hyvinkin mielenkiintoinen kehityskohde olisi pyrkiä mallintamaan sitä kuinka Loton potti inflatoituu ajassa, siihen ei tämän työn puitteissa pystytty. Lotto uudistui 41/2011 ja uuden pelin dataa ei hyödynnetty analyseissä. Oletan että Loton pelaamiseen vaikuttavat tekijät muuttuvat suhteellisen hitaasti ajassa, mutta siitä huolimatta voisi olla mielenkiintoista tutkia onko pelaamiseen vaikuttavat tekijät ja etenkin niiden vaikuttavuuden suuruusluokka merkittävästi muuttuneet uudistuneen pelin myötä.

Katson tämän työn opettaneen minulle erittäin paljon siitä, mitkä tekijät Loton pelaamiseen vaikuttavat ja kuinka Loton vuositason myyntien voidaan olettaa käyttäytyvän pitkässä juoksussa. Työn suurin anti ei niinkään ole itse tässä diplomityössä vaan siinä, että koen työtä tehdessä oppineeni paljon menetelmiä ja tietotaitoa, joita voin tehokkaasti hyödyntää tulevaisuuden työelämän haasteissa.

LÄHDELUETTELO

Pindyck, Robert S and Rubinfeld, Daniel L. Econometric Models & Economic Forecasts. (1991). McGRAW-HILL INTERNATIONAL EDITIONS. third edition.

Mellin, I. "Lineaarinen regressioanalyysi" & "Yleinen lineaarinen malli (2008)" From <https://noppa.aalto.fi/noppa/app>--Teknillinen Korkeakoulu.
<https://noppa.aalto.fi/noppa/kurssi/mat-2.3128/materiaali>. Haettu 1.6.2012

Mellin, I. Todennäköisyys: Kaavat. (2007).

Ylikangas, H. Markat Miljooniksi, puoli vuosisataa suomalaista veikkausta. (1993) Gummerus Kirjapaino Oy.

Bernard, B., Bruneel, C., Houtman De Smedt, H., Maffesoli, M., Muchembled, R., Van Eenoo, R. (1994) Lotteries in Europe: Five centuries of history. Belgian National Lottery.

Kupiainen, K. Lottopelin lanseeraus Suomeen. (1971).

Järvinen, H. Rahapeliä toiminnallinen ja matemaattinen rakenne - pelinjärjestäjän näkökulma. (2007).

Valtioneuvoston asetus Veikkaus Oy:n rahapeliä toimeenpanosta. (1288/2011). <http://www.finlex.fi/fi/laki/alkup/2011/20111288>.

Arpajaislaki 23.11.2001/1047. <http://www.finlex.fi/fi/laki/ajantasa/2001/20011047>.

Veikkaus Oy:n rahapeliä säännöt. http://www.finlex.fi/data/normit/38381-VeikkausOyRahapeliäSaannotSMDno2011_2027_010112.pdf.

Forrest, D., Simmons, R., Chesters, N. (2002) Buying a dream: Alternative models of demand for lotto. Economic Inquiry, Vol. 40, No. 3, July 2002, 485-496.

Walker, I., Young, J. (2001) An Economist's Guide to Lottery Design. The Economic Journal, Vol. 111, No. 475, Nov., 2001, pp. F700-F722.

Beenstock, M., Goldin, E., Haitovsky, Y. (200) What jackpot? The optimal Lottery tax. European Journal of Political Economy, Vol. 16, (2000), 655-671

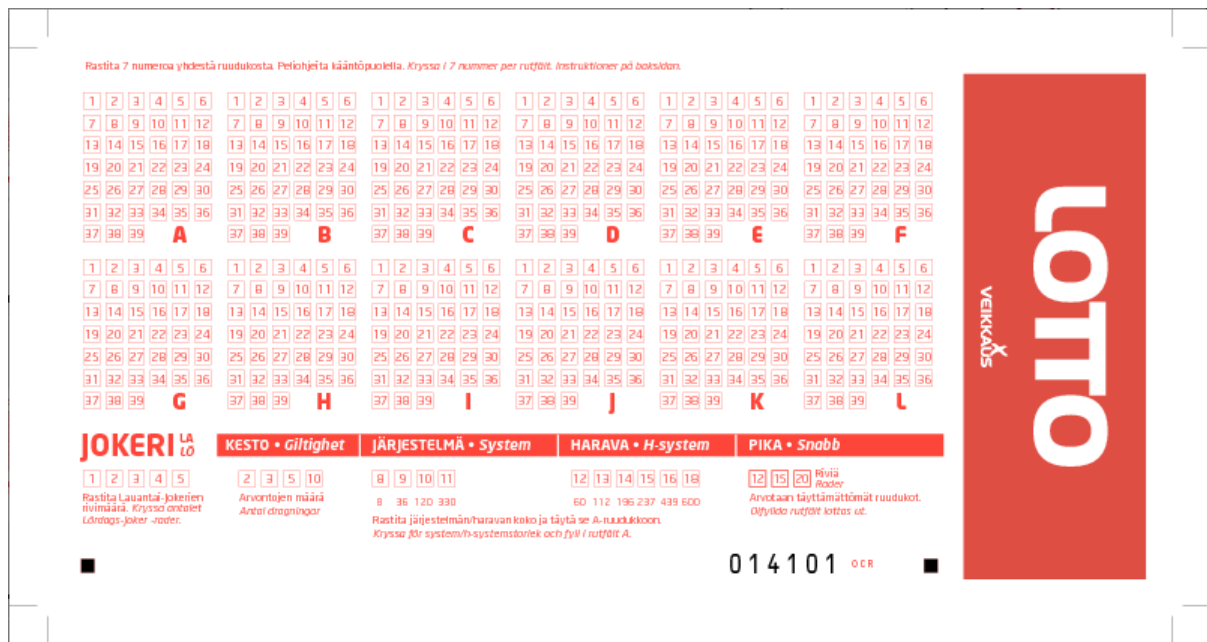
Koenker, R. 1981. (1981) A note on studentizing a test for heteroscedasticity, Journal of Econometrics, Elsevier, vol. 17(1), pages 107-112, September.

Jarque, C. M., Bera, A. K. (1987) A test for normality of observations and regression residuals. International Statistical Review. Vol. 55, No. 2, 1987, pp. 163-172.

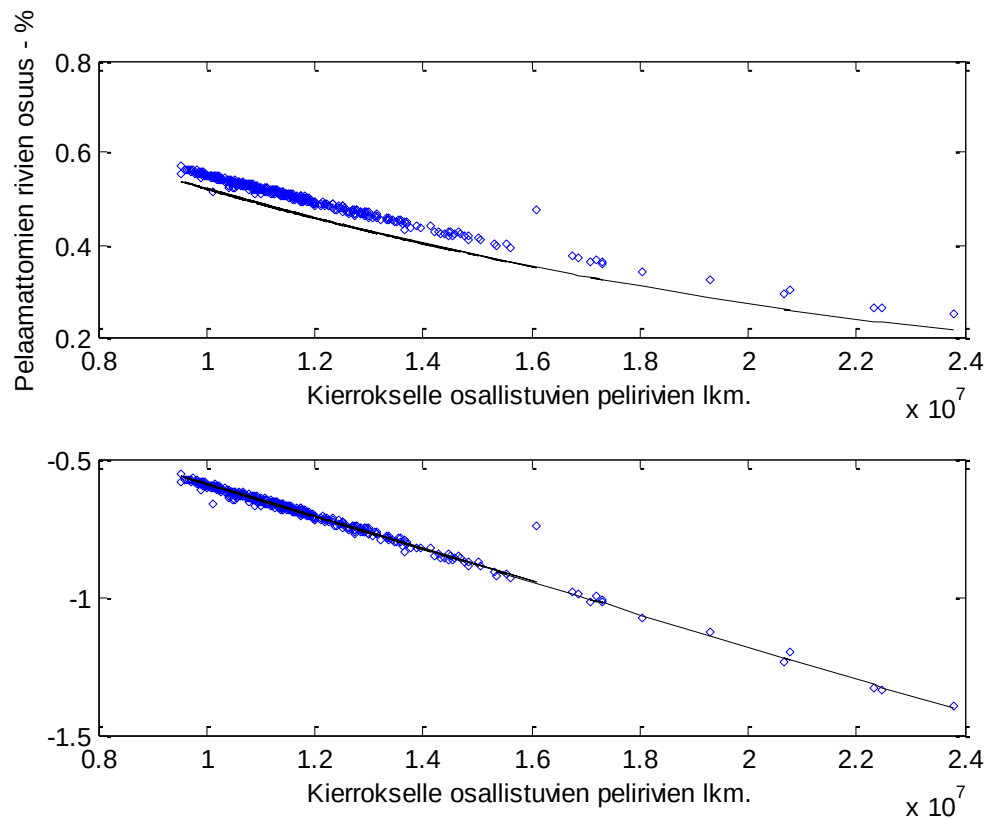
Procedures for Detecting Outlying Observations in Samples

Grubbs, F .E., (1969) Procedures for Detecting Outlying Observations in Samples.
Technometrics, Vol. 11, No. 1 (Feb., 1969), pp. 1-21

LIITTEET



Kuva 19: Esimerkki Lottokupongin etuosasta (käyttöönotto 41/2011)



Kuva 20: Pelikäyttäytymisen vaikutus ruudukoiden täyttöasteeseen

Loton säännöt

Voimaan 1.1.2012

Lottoon sovelletaan Veikkaus Oy:n rahapeliä toimeenpanosta annetun valtioneuvoston asetuksen ja Veikkaus Oy:n rahapeliä pelisäännöistä annetun sisäasiainministeriön asetuksen mukaan seuraavaa:

I YLEISTÄ

1. Lotto on numeroveikkaus, jossa arvotaan seitsemän (7) numeroa ja kaksi (2) lisänumeroa kolmestakymmenestäyhdeksästä (39) numerosta.
2. Lottoon osallistutaan Veikkaus Oy:n (yhtiön) pelijärjestelmän välityksellä. Pelitiedot annetaan yhtiön pelijärjestelmään joko asiamiespisteissä tai itsepalveluna puhelimen, tietokoneen taikka niihin verrattavan teknisen sovelluksen välityksellä.
3. Pelaajan ja yhtiön välinen pelisopimus tulee voimaan, kun pelimaksu on maksettu ja yhtiön pelijärjestelmässä olevat pelitiedot on siirretty (yhtiön pelijärjestelmästä) virallisten valvojen tietojärjestelmään.
4. Yhtiön pelijärjestelmään on määritelty pelaamismuodot, pelien voimassaoloajat ja osallistumismaksut.
5. Pelaaminen voi tapahtua pelaajan ilmoittamalla tai yhtiön pelijärjestelmän arpomilla pelitiedoilla niin sanottuna pikapelinä taikka pelitiedot sisältävällä valmiskupongilla.
6. Lottoa voi myös pelata järjestelmä- tai haravapelinä.
Järjestelmä on pelaamismuoto, jossa samassa pelissä on useampi kuin yksi pelirivi. Järjestelmän enimmäiskoko on määritelty yhtiön pelijärjestelmässä.
Harava on pelaamismuoto, jossa samassa pelissä on useampi kuin yksi pelirivi. Siinä on vähemmän pelirivejä kuin vastaavassa järjestelmäpelissä. Haravaan osallistuvat pelirivit ja haravien koot on määritelty yhtiön pelijärjestelmässä.
Jokainen pelirivi, joka sisältyy järjestelmään tai haravaan, on itsenäinen veikkaus.
7. Pelaaja voi osallistua yhdestä kolmeen (1 - 3), viiteen (5) tai kymmeneen (10) perättäiseen arvontaan. Lottoa voi pelata itsepalveluna myös toistaiseksi voimassa olevana niin sanottuna ikipelinä.
8. Yhtiön pelijärjestelmä tulostaa pelaajalle pelijärjestelmään tallentuneet pelitiedot sisältävän pelitositteen.
9. Pelitositteesta ilmenee arvonta tai arvonnat, joihin peli osallistuu.
10. Pelaajan on tarkistettava yhtiön pelijärjestelmän tulostaman pelitositteen sisältö ja esitettävä sitä koskevat huomautukset välittömästi pelaamisen tapahduttua.
11. Itsepalveluna pelattaessa pelaajalle ei tulostu pelitositetta. Itsepalveluna pelattaessa pelaaja hyväksyy pelin ennen sen siirtämistä yhtiön pelijärjestelmään. Pelin muuttaminen tämän jälkeen ei ole mahdollista.

12. Itsepalveluna pelattaessa pelaajalla on mahdollisuus saada yhden vuoden ajan pelaamisesta erikseen pyynnöstä tieto peleistään, jotka ovat tallentuneet yhtiön pelijärjestelmään.

13. Itsepalveluna pelattaessa peli hyväksytään yhtiön pelijärjestelmään edellyttäen, että pelimaksu on suoritettu pelaajan pelitililtä tai pankkitililtä.

14. Pelitiedot siirretään yhtiön pelijärjestelmästä virallisten valvojen tietojärjestelmään.

15. Pelisopimus tulee kuitenkin voimaan vain, jos peli on tallentunut yhtiön pelijärjestelmään ja virallisten valvojen tietojärjestelmään ennen kuin arvonta on alkanut. Yhtiö palauttaa pelipanoksen pelisopimuksesta, joka ei ole tullut voimaan. Pelipanos palautetaan joko itsepalveluna pelatuista peleistä pelaajan yhtiössä olevalle pelitilille, pelaajan pelaamisen yhteydessä ilmoittamalle pankki- tai pelitilille tai pyynnöstä yhtiön pelijärjestelmästä tulostettua pelitositetta vastaan.

16. Yhtiöllä on oikeus riskienhallinnallisista syistä kieltäytyä vastaanottamasta kokonaan tai osittain pelejä.

II ARVONTA

17. Loton arvonta suoritetaan pelaajan päätyttyä Poliisihallituksen hyväksymällä ja virallisten valvojen hallussa olevalla arvontalaitteistolla.

18. Jos arvontaa ei voida toteuttaa arvontalaitteistolla, se suoritetaan virallisten valvojen valvonnassa käsin nostoarvontana.

19. Jos arvonta keskeytyy, arvotut numerot jäävät voimaan. Numero katsotaan arvotuksi, kun se on siirtynyt kokonaan arvontarummun ulkopuolelle.

Jos arvontalaitteistoon tullutta teknistä vikaa ei voida viipymättä korjata, loput numerot arvotaan virallisen valvojan valvonnassa käsin nostoarvontana.

III VOITTOLUOKAT JA VOITTO-OSUUDET

20. Loton voittoluokat ovat:

1. voittoluokka seitsemän (7) oikein.
2. voittoluokka kuusi (6) ja yksi lisänumero oikein.
3. voittoluokka kuusi (6) oikein.
4. voittoluokka viisi (5) ja kaksi lisänumeroa oikein.
5. voittoluokka viisi (5) ja yksi lisänumero oikein.
6. voittoluokka viisi (5) oikein.
7. voittoluokka neljä (4) ja kaksi lisänumeroa oikein.
8. voittoluokka neljä (4) ja yksi lisänumero oikein.
9. voittoluokka neljä (4) oikein.
10. voittoluokka kolme (3) ja kaksi lisänumeroa oikein.
11. voittoluokka kolme (3) ja yksi lisänumero oikein.

Yksittäisessä pelirivissä voi olla vain yksi voittotulos.

Loton päävoitto-osuus maksetaan korotettuna yhtiön ennalta määräämillä kierroksilla niille Loton päävoittoon oikeutetuille, joiden varsinaiseen arvontaan osallistuneella päävoittotositteella pelatulla jokerinumerolla on voitettu mikä tahansa Lauantai-Jokerin voitto, samalla kierroksella millä päävoittotosite on oikeutettu Loton korotet-

tuun päävoittoon. Yhtiön on ilmoitettava kirjallisesti 14 vuorokautta etukäteen Poliisihallitukselle, millä pelikierroksilla korotus on voimassa ja minkä suuruinen korotus on.

21. Loton kierroskohtaisesta kokonaisvaihdosta laskettu jakosumma on 41,1 prosenttia.

22. Loton kierroskohtaisesta kokonaisvaihdosta 16,5 prosenttia jaetaan tasan ensimmäisen voittoluokan tulosten kesken. Loton kierroskohtaisesta kokonaisvaihdosta siirretään 2,0 prosenttia maksettavaksi pelin myöhemmillä kierroksilla yhtiön määrämällä tavalla.

23. Loton kierroskohtaisesta kokonaisvaihdosta lasketusta jakosummasta maksetaan seuraavat kiinteät voitot:

kolme (3) ja kaksi lisänumeroa oikein (10. voittoluokka)	5,00 euroa
kolme (3) ja yksi lisänumero oikein (11. voittoluokka)	1,00 euro

24. Loton kierroskohtaisesta kokonaisvaihdosta lasketusta jakosummasta jaetaan seuraavat osuudet kiinteiden voittojen ja ylimmän voittoluokan jako-osuuden sekä myöhemmille kierroksille siirrettävän osuuden jälkeen:

kuusi (6) ja yksi lisänumero oikein (2. voittoluokka)	8,8 %
kuusi (6) oikein (3. voittoluokka)	10,0 %
viisi (5) ja kaksi lisänumeroa oikein (4. voittoluokka)	2,0 %
viisi (5) ja yksi lisänumero oikein (5. voittoluokka)	4,6 %
viisi (5) oikein (6. voittoluokka)	13,5 %
neljä (4) ja kaksi lisänumeroa oikein (7. voittoluokka)	2,7 %
neljä (4) ja yksi lisänumero oikein (8. voittoluokka)	16,8 %
neljä (4) oikein (9. voittoluokka)	41,6 %

25. Jos ensimmäisessä voittoluokassa ei löydy voittoon oikeuttavaa tulosta, tämän voittoluokan kierroskohtaisesta kokonaisvaihdosta laskettu jakosumma siirtyy jaettavaksi pelin myöhemmin ilmoitettaville kierroksille yhtiön määrämällä tavalla.

26. Jos jollakin kierroksella ei ole yhtään "6 ja yksi lisänumero oikein" tulosta (2. Voittoluokka), toisen voittoluokan kierroskohtaisesta kokonaisvaihdosta laskettu jakosumma siirretään kokonaisuudessaan seuraavalle kierrokselle jaettavaksi "6 ja yksi lisänumero oikein" tulosten kesken.

27. Jos seuraavalla tai sitä seuraavillakaan kierroksilla ei ole yhtään "6 ja yksi lisänumero oikein" tulosta (2. Voittoluokka), toisen voittoluokan kierroskohtaisesta kokonaisvaihdosta laskettu jakosumma sekä edellisillä kierroksilla jakamatta jääneet toisen voittoluokan kierroskohtaisesta kokonaisvaihdosta lasketut jakosumat siirretään kokonaisuudessaan edelleen seuraavalle kierrokselle jaettavaksi "6 ja yksi lisänumero oikein" tulosten kesken.

28. Jos jollakin kierroksella ei ole yhtään tulosta jossain voittoluokista 3. - 9., jaetaan kyseisen voittoluokan voittoina jaettava osuus tasan niiden voittoluokkien 3. - 9. kesken joissa tulos kierroksella löytyy. Voittoluokkien 3. - 9. voitto-osuutena maksetaan aina vähintään 5,00 euroa.

29. Ensimmäiseen ja/tai toiseen voittoluokkaan voidaan vahvistaa kierroskohtaisesta kokonaisvaihdosta riippumaton vähimmäisjakosumma.

IV VOITONMAKSU

30. Voittojen maksu alkaa arvontaa seuraavana päivänä. Voitto, joka on vähintään kaksikymmentätuhatta (20 000) euroa maksetaan pelaajalle aikaisintaan kolmen viikon kuluttua oikean tuloksen vahvistamisesta.

31. Voitto maksetaan ainoastaan pelitositetta vastaan, jollei pelaajan pankkiyhteystietoa tai pelitilitietoa ole tallennettu yhtiön pelijärjestelmään pelaamisen yhteydessä. Pelitosite on voittoa lunastettaessa luovutettava yhtiölle tai sen asiamiehelle.

32. Voittajan henkilöllisyys on todennettava tuhat (1000) euroa ja sitä suurempia voittoja lunastettaessa, jollei pelaajan pankkiyhteystieto tai pelitilitieto ole tallentunut yhtiön pelijärjestelmään pelaamisen yhteydessä.

33. Itsepalveluna pelattaessa voitto maksetaan pankkitilille tai pelaajan yhtiössä olevalle pelitilille.

34. Pelitosite on mitätön, jos sen luettavuus on huonontunut niin, ettei sen sisältämä pelitapahtuma ole yksilöitävissä.

35. Jos pelitositteessa, yhtiön pelijärjestelmässä tai virallisten valvojien tietojärjestelmässä olevat pelitiedot eroavat, voimassa ovat virallisten valvojien tietojärjestelmään tallentuneet tiedot.

36. Loton voitot on perittävä yhden vuoden kuluessa pelisääntöjen mukaisen lopputuloksen vahvistamisesta. Pelaaja menettää oikeuden voittoon, jota hän ei ole nostanut määräajassa.