M. Palo

SYSTEMAATTISESTA RYVÄSOTANNASTA JA SEN KÄYTÖSTÄ HAKKUUPOISTUMAN MÄÄRITYKSESSÄ

Risto Seppälä

SISÄLLYSLUETTELO:

H CO_2, H_2O + ravinteet sekä auringon valo ?

1. JOHDANTO

Jotta päästäisiin kestävään ja tasaiseen tai mieluummin vähitellen kohoavaan puun tuotokseen metsätalous tarvitsee suunnittelua. Kaiken suunnittelun lähtökohtana on vallitsevan tilanteen tunteminen. Metsää kokonaisuudessaan voidaan tarkastella jatkuvaprosessisena systeeminä, jonka syötön muodostaa val ~~metsän kasvu~~ ja tulostuksen puun poistuma. Näistä molemmista tekijöistä, joiden avulla voidaan seuraten metsävarannon kehitystä ~~muodostaa käsitys metsätalouden kunkinhetkisestä tilasta~~, tarvitaan jatkuvasti luotettavia ja ajan tasalla olevia tietoja.

Puuston kasvusta tarvittavan informaation hankkimisessa käytetään erilaisia metsänarvioimismenetelmiä. Inventoinnilla selvitetään kasvun lisäksi yleensä myös tutkittavan alueen puuston määrä ja rakenne, osittain hakkuumäärä sekä erilaiset pinta-aloja koskevat jakaumat. Näiden menetelmien perustana on Suomessa pitkään ollut linja-arviointi, joka juontaa juurensa jo viime vuosisadan jälkipuoliskolta. Siitä ollaan kuitenkin vähitellen luopumassa, ja arviointimenetelmien kehitys vaikuttaakin tällä hetkellä hyvin ekspansiiviselta.

Metsän kasvu muodostaa perustan metsän tuotokselle. Puun kokonaispoistuman määrittäminen tiettynä aikavälinä ei voi tapahtua pelkästään inventoimalla puuston määrä kahtena mittaushetkenä. Mikäli olisi tarkasti selvitettävissä puuston kasvu mittaushetkien välisenä aikana, saataisiin puustosta tapahtunut poistuma lasketuksi lisäämällä kasvu inventointiajankohtien puuston määrien erotukseen.

Poistuman määritys on kuitenkin Suomessa tapahtunut yleensä ~~suurelta osin~~ erillisenä tutkimuksena. ~~Pääpiirteissään~~ voidaan Kokonaispoistuma jakaa hakkuupoistumaan ja luonnonpoistumaan. Viimeksi mainittu, se ~~joka~~ on suuruudeltaan noin kaksi prosenttia kokonaispoistumasta, perustuu vain arviolukuihin. Tämän vuoksi tutkimuksen kohteena on (yleensä) ollut hakkuupoistuman määrittäminen. Itse asiassa olisi ehkä parempi puhua raakapuun tuotoksen määrittämisestä, jolloin hakkuupoistumaan on lisättävä luonnonpoistumasta

H tule

käyttöön tullut osa.

Kaikkein lähimmäksi itse puuta tuottavaa systeemiä on menty kantomittausmenetelmässä, jossa mittaus kohdistetaan hakattujen puiden kantoihin. Ideana kantomittaus on peräisin jo tämän vuosisadan alusta. Tällä hetkellä perustuvat Ruotsin hakkuumääräarviot tähän menetelmään, jota on yhtäjaksoisesti käytetty jo viidentoista vuoden ajan. Suomessa suoritetuissa kokeiluissa on havaittu menetelmän tuottavan 5-10 % todellista poistumaa pienempiä arvoja.

Hakkuupoistuman määrittämistä puun tuottajaan eli metsänomistajaan kohdistuvin näyttein on myös kokeiltu. Suoritetuissa tutkimuksissa on hakkuutietoja koskeva aineisto kerätty metsälökohtaisin haastatteluin. Pahimpina ongelmina ovat olleet tutkittavan perusjoukon luotettava luettelointi ja haastattelupisteiden hajanaisuudesta johtuvat suuret matkakustannukset. Tällä menetelmällä saadut alustavat tulokset vaikuttavat melko hyviltä. Tämä edellyttää kuitenkin, että arviointien suorittaminen pääasiassa haastattelun avulla ei aiheuta systemaattista virhettä.

Raakapuun ostajat ovat eräs ryhmä, johon puun tuotoksen tutkiminen voidaan kohdistaa. Kustannuksiltaan tämä tapa tulee halvemmaksi kuin metsänomistajiin kohdistuva tiedustelu, koska ostajien lukumäärä on huomattavasti omistajie määrää pienempi. Suurimpana heikkoutena tässä menetelmässä on se, että kirjaamatta jäävät kokonaan kiinteistöjen omasta metsästä omaan käyttöön hankkima puu sekä metsänomistajien suoraan yksityiselle kotitarveostajalle m ymä puu.

Kaikkein kauimmaksi puun tuotantosysteemistä mennään johdettaessa poistuma käytön perusteella. Tällaista menetelmää sovelletaan Suomessa tällä hetkellä. Kiinteistöjen käyttämää puuta lukuunottamatta saadaan lähes kaikki muu puunkäyttö selvitetyksi vuosittain. Otantaan perustuvat kiinteistöjen puunkäytön tutkimukset sen sijaan on tehty noin 10 vuoden välein, ja niistä viimeinen kohdistui vuosiin 1964-1966. Kaikkiaan muodostaa kiinteistöjen käyttämä puu yli 20 % koko hakkuupoistumasta. Sen osuutta ei siis ~~voi~~ aliarvioida, vaikka se sekä määrällisesti että suhteellisesti onkin pienenemässä.

Kaikkiin edellä esitettyihin menetelmiin ~~on~~ sisältyy puutteita. Onneksi ne eivät kuitenkaan kaikki ole päällekkäisiä. Niinpä tulee ~~helposti~~ mieleen ajatus ottaa eri menetelmien hyviä puolia ja niitä yhdistämällä luoda entistä tehokkampi kokonaisniin puiden kasvupaikkaa ?

menetelmä. Tämän kirjoittajan näkemyksen mukaan suuntaa-antavana linjana menetelmän etsimisessä tulee olla pyrkimys päästä mahdollisimman lähelle metsää. Vahvan rajoituksen tälle pyrkimykselle kuitenkin asettavat kustannukset, joilla on taipumus suurentua sen mukaan, mitä lähemmäksi tutkimuskohdetta edetään.

Koska tietoja kokonaispoistumasta tarvitaan vuosittain, on pääosa tiedoista saatava yhtä useasti. Kirjoittajalle muotoutuneen käsityksen mukaan tähän soveltuu parhaiten raakapuun ostajiin kohdistuva tiedustelu. Suurimpien ostajien lukumäärä on niin pieni, että niiden osalta tiedustelu voidaan tehdä kokonaisnäytteenä. Pienimmistä ostajista sen sijaan on otettava otos, mutta vuosittainkaan suoritettuna se ei muodostune kohtuuttoman kalliiksi. Vaikeimman ongelman muodostaa ilmeisesti se, että perusjoukko muuttuu jatkuvasti, joten otantakehikon pitäminen ajan tasalla vaatii ponnistuksia. Pysyvän asiamiesverkoston avulla se ei kuitenkaan ole ylivoimainen tehtävä.

Kuten aikaisemmin todettiin, jää ostajiin perustuvassa tutkimuksessa osa hakkuupoistuman eristä kirjaamatta. Tämä osa koostuu siitä puusta, joka käytetään ostajien ryhmään kuulumattoman puuntuottajan omiin tarkoituksiin tai myydään siten, ettei se tule kirjatuksi raakapuun ostajien toimesta. Näiden erien selvittäminen metsänomistajien avulla tuntuu luonnolliselta. Koska metsälöiden lukumäärä Suomessa on lähes 400 000, ~~ei voi ajatellakaan koko perusjoukon tutkimista, vaan~~ estimointi on suoritettava otosta käyttäen. Samalla kun selvitetään puuttuvan osan suuruus, voidaan tutkia myös koko hakkuupoistuma ja näin saada vertailuaineistoa ostajatutkimuksen rinnalle. Kirjoittajan esitys käytettävästä otantamenetelmästä on seuraava.

Menetelmän perustana on ryväsotanta. Syyt tähän löytyvät kahdelta suunnalta. Ensinnäkin metsälöiden suuri lukumäärä tekee tavalliseen satunnaisotantaan perustuvan otantakehikon luomisen työlääksi. Jos konstruoidaan ryväs sopivasti, tarvitaan informaatiota vain otosrypäistä. Toisena syynä ovat haastatteluun liittyvät matkakustannukset. Tavanomaista yksiasteista satunnaisotantaa käytettäessä hajaantuvat otosmetsälöt niin kauas toisistaan, että nykyisen hyvän tieverkoston aikakaudellakin aiheuttavat matkat metsälöstä toiseen huomattavan suuria kustannuksia. Niitä lisää vielä se, että omistajan ollessa poissa kotoa joudutaan käynti uusimaan.

Rypäiden poimimisessa otokseen on ollut tarjolla kaksi var-

H-vuotta harvemmin aikavälein.

teenotettavaa vaihtoehtoa: yksinkertainen satunnaisotanta ja systemaattinen otanta. Viimeksi mainitulla on esillä olevan ryväsotannan kannalta niin olennaisia käytännöllisiä etuja, että valinnassa on päädytty siihen.

Koska tällainen puun tuottajiin kohdistuva otanta tulee vaatimaan suhteellisesti huomattavasti enemmän kustannuksia kuin ostajanäytteen avulla mitatun erän estimointi, on tarkoituksenmukaista tyytyä suorittamaan se ~~pitemmin kuin vuoden aikavälein~~. Puunkäyttötutkimuksissa sovellettu 10 vuotta tuntuu liian pitkältä ajanjaksolta, jotta välivuosien inter- ja ekstrapolointi pystytään suorittamaan riittävän luotettavasti. Sen vuoksi siirtyminen esimerkiksi puolta lyhyempään aikaan tuo jo mukanaan huomattavan parannuksen.

Koska tutkimusta toisen ja sitä useamman kerran käynnistettäessä tulee olemaan tietoja aikaisemmin saaduista estimaateista, voidaan niitä muutosten estimoinnin lisäksi käyttää hyväksi parantamaan tulosten tarkkuutta. Tämän tehokas toteuttaminen edellyttää suhde- tai regressioestimaattien käyttöä ja otosyksiköiden ainakin osittaista pysyttämistä samoina ajanjaksosta toiseen. Lopuksi tarkastellaankin ketjutusratkaisuun perustuvaa estimointia.

Päälinjana seuraavassa tulee olemaan se, että ensin selvitetään yleisesti otantamenetelmä kulloisellakin tasolla ja sen jälleen sovelletaan saatuja tuloksia ongelmakentän muodostavaan hakkuupoistuman määritykseen. Käytettävät numeroarvot ovat pääosiltaan ~~suuressa määrin~~ hypoteettisia, koska todellisia lukuja ei olleet ~~ole~~ saatavissa.

2. RYVASOTANTA

2.1. Ryväsotannan pääperiaatteet

Yleisesti hyväksytyt periaatteet otantamenetelmän valintaa suoritettaessa ovat joko tulosten tarkkuuden maksimointi annettuihin kustannuksiin nähden tai kustannusten minimointi annettuun tarkkuusvaatimukseennähden. Maantieteellisesti hajallaan olevia otantayksiköitä mitattaessa nousevat kustannukset yksiasteista satunnaisotantaa käytettäessä huomattav~~iksi~~ an korkeiksi. Usein on lisäksi otantakehikon luominen perusjoukon vaillinaisen luetteloinnin takia työlästä ja kallista, monesti jopa mahdotontakin. Tällaisissa tapauksissa tarjoaa ryväsotanta erään mahdollisuuden saavuttaa käyttötarkoituksia silmällä pitäen riittävän tarkkoja tuloksia kohtuullisin kustannuksin.

Estimoitavan muuttujan mittaus kohdistuu ryväsotannassa sen alkeisyksikköön. Kun rypäälle lasketaan keskilukuja, ne ilmaistaan alkeisyksikön tasolla. Sen määrääminen riippuu ennen kaikkea tutkimuksen tarkoituksesta. Useissa tapauksissa samassa tutkimuksessa voi olla useampia kuin yksi alkeisyksikkö.

Ryväsotannassa jaetaan perusjoukko jonkin ominaisuuden, esimerkiksi maantieteellisen sijainnin perusteella, ryhmiin, jotka toimivat otannan yksikköinä. Kun tällainen ryväs on valittu näytteeseen, voidaan mitata tulokset kaikista siihen sisältyvistä alkeisyksiköistä tai valita vain osa niistä mittauksen kohteeksi. Jälkimmäisessä tapauksessa on kyse moniasteotannasta.

Milloin tiettyjä maantieteellisiä kokonaisuuksia käytetään rypäinä, puhutaan alueotannasta. Alueen voi muodostaa esimerkiksi kaupunginosa, metsälö tai maastoon merkitty koeala. Alueotanta saattaa monessa tapauksessa tarjota ainoan mahdollisuuden otantakehikon luomiseen. Tällaiseen alueeseen sisältyvistä alkeisyksiköistä ei rypään täydellisessä mittauksessa tarvitse olla välttämättä lainkaan etukäteisinformaatiota. Tärkeätä on tällöin vain, että kukin yksikkö on yksiselitteisesti luettavissa kuuluvaksi vain yhteen rypääseen.[1)]

---

1) Hansen, Hurwitz and Madow: Sample Survey Methods and Theory, Vol. I, New York 1962, s. 246-247.

2.2. Ryväsotannan tunnuslukujen laskeminen

Seuraavassa rajoitutaan siihen käytännössä kaikkein yleisimpään tapaukseen, jossa rypäät eivät alkeisyksikköjensä lukumäärän suhteen ole yhtä suuria. Otosrypäiden valinnassa oletetaan käytetyn osittamatonta satunnaisotantaa, jossa kaikille perusjoukon rypäille on annettu yhtä suuri todennäköisyys joutua näytteeseen.

Käsittäköön i:s ryväs M_i elementtiä ja olkoon x tällaisesta elementistä mitattava suure. Silloin on i:nnen rypään keskiarvo

$$(2.1) \qquad \bar{x}_i = \frac{1}{M_i} \sum_{j=1}^{M_i} x_{ij} .$$

Kaikille otokseen sisältyville rypäille voidaan laskea estimaatti (2.1). Mikäli kaikkia rypään elementtejä ei oteta mukaan, vaan otos tehdään kaksiasteisena, korvataan M_i otoselementtien lukumäärällä m_i.

Usean rypään otoksesta voidaan elementtiä kohti laskettu keskiarvo saada monella tavalla. Seuraavassa esitetään kolme eri menetelmää. Kunkin estimaattorin kohdalla käydään läpi sekä yksi- että kaksiasteisen otannan tapaukset.[1]

Yksinkertaisin estimaatti saadaan ottamalla ryväskeskiarvojen aritmeettinen keskiarvo

$$(2.2) \qquad \bar{\bar{x}} = \frac{1}{n} \sum_{i=1}^{n} \bar{x}_i .$$

n = otosrypäiden luku

Tämä estimaatti antaa harhaisen tuloksen, koska kaikilla perusjoukon elementeillä ei ole ollut yhtä suurta todennäköisyyttä joutua mukaan näytteeseen. Harhan vaikutus ei kuitenkaan ole huomattava, jos otosrypäiden määrä on suuri ja alkeisyksiköiden luku eri rypäissä ei vaihtele kovin paljon. Mikäli rypäät olisi valittu niiden kokoon nähden vaihtelevin todennäköisyyksin, olisi tämä estimaatti harhaton.

Keskiarvon (2.2) varianssi on yksiasteotannan tapauksessa

---

1) Esitys perustuu pääosiltaan teokseen Sukhatme: Sampling Theory of Surveys With Applications, Iowa 1963, s. 265-69 ja s. 315-35.

normaali yksinkertaisen satunnaisotannan varianssi, jossa elementtien välisen vaihtelun sijasta estimoidaan rypäitten välistä vaihtelua.

$$(2.3) \qquad V(\bar{x}) = (1 - \frac{n}{N}) \frac{s_b^2}{n} = (1 - \frac{n}{N}) \frac{\sum_{i=1}^{n} (\bar{x}_i - \bar{x})^2}{n(n - 1)} .$$

Kaksiasteotannan tapauksessa eivät yksityisen rypään kaikki elementit ole mukana, joten varianssissa täytyy estimoida myös rypäiden sisäinen vaihtelu. Keskiarvon varianssi saa tällöin muodon

$$(2.4) \qquad V'(\bar{x}) = (1 - \frac{n}{N}) \frac{s_b^2}{n} + \frac{1}{nN} \sum_{i=1}^{n} (1 - \frac{m_i}{M_i}) \frac{s_i^2}{m_i} .$$

Vasemmanpuoleinen termi koostuu rypäiden välisestä ja oikeanpuoleinen niiden sisäisestä vaihtelusta. Termi s_i^2 lasketaan normaalilla tavalla rypäässä olevien havaintojen ja niiden keskiarvon poikkeamien suhteen. Jos perusjoukon rypäiden lukumäärä on suuri, lähenee jälkimmäinen termi nollaa ja edellisen termin äärellisen populaation korjauskerroin ykköstä, jolloin varianssille saadaan aproksimaatio kaavasta s_b^2 / n . Monissa käytännön tutkimuksissa tämä likiarvo antaa täysin tyydyttävän tuloksen.

Perusjoukon elementtien keskiarvolle voidaan muodostaa myös harhaton estimaatti. Yksinkertaisin niistä on muotoa

$$(2.5) \qquad \bar{x}' = \frac{1}{n} \sum_{i=1}^{n} \frac{M_i}{\bar{M}} \bar{x}_i , \text{ jossa } \qquad \bar{M} = \sum_{i=1}^{N} M_i / N$$

eli $\bar{M}$ on perusjoukon rypäisiin sisältyvien elementtien keskimääräinen luku. Keskiarvoestimaatin harhattomuus on helposti nähtävissä, sillä

$$E(\bar{x}') = \frac{1}{n\bar{M}} \sum_{i}^{n} E(M_i \bar{x}_i) = \frac{1}{n\bar{M}} \frac{n}{N} \sum_{i}^{N} M_i \bar{x}_i = \bar{X} \text{ , kun}$$

kaikilla perusjoukon yksiköillä on ollut sama todennäköisyys n/N joutua otokseen.

Otosvarianssi voidaan yksiasteotannan tapauksessa kirjoittaa seuraavasti:

$$(2.6) \qquad V(\bar{x}') = (1 - \frac{n}{N}) \frac{s_{b'}^2}{n}, \text{ jossa}$$

$$s_{b'}^2 = \frac{1}{n-1} \sum_i^n (\frac{M_i}{\bar{M}} \bar{x}_i - \frac{1}{n} \sum_i^n \frac{M_i}{\bar{M}} \bar{x}_i)^2$$

$$= \frac{1}{n-1} \sum_i (\frac{M_i}{\bar{M}} \bar{x}_i - \bar{x}')^2 .$$

Kaavasta voi helposti nähdä, että varianssin suuruus riippuu tulon $M_i\bar{x}_i$ vaihtelusta. Estimaattori (2.6) antaakin korkeampia arvoja kuin estimaattori (2.3), jos rypään koon ja sen keskiarvon välillä vallitsee positiivinen korrelaatio.

Kaksiasteotantaa vastaavaan varianssiin tulee jälleen mukaan myös rypäiden sisäinen vaihtelu.

$$(2.7) \qquad V'(\bar{x}') = (1 - \frac{n}{N}) \frac{s_{b'}^2}{n} + \frac{1}{nN} \sum_i (\frac{M_i}{\bar{M}})^2 (1 - \frac{m_i}{M_i}) \frac{s_i^2}{m_i} .$$

N:n ollessa suuri saadaan likiarvo samalla tavalla kuin estimaattorille (2.4).

Käytännön metsätieteellisissä tutkimuksissa on käynyt ilmi, viite? että estimaattorin (2.2) antamat tulokset ovat usein pahasti harhaisia, koska elementtien lukumäärät vaihtelevat huomattavasti rypäästä toiseen. Estimaattorin (2.5) varianssi taas muodostuu tavallisesti suureksi, ja lisäksi sen käyttö vaatii tietoja koko perusjoukon rypäistä. Tämän vuoksi onkin osoittautunut hyväksi käyttää seuraavaa estimaattoria, jolla laskettaessa ei tarvitse olla selvillä perusjoukon elementtien kokonaismäärästä. Keskiarvoestimaatti muodostetaan painottamalla ryväsarvoja rypääseen sisältyvien alkeisyksikköjen kokonaismäärällä. Otoksen aikaan saanut prosessi on ollut tällöin sellainen, että jokaiselle perusjoukon otosyksikölle x_{ij} on annettu sama todennäköisyys n/N tulla valituksi kyseiseen otokseen.

$$(2.8) \qquad \bar{x}'' = \frac{\sum_i^n M_i \bar{x}_i}{\sum_i M_i} .$$

Itse asiassa tämä estimaattori on luonteeltaan suhde-estimaattori, koska sekä osoittajassa että nimittäjässä olevat termit ovat satunnaismuuttujia. Kaavan avulla saatava keskiarvo on harhainen, mutta tarkentuva otosrypäitä lisättäessä. Harhan vaikutus pienenee myös, jos rypään koko ja sen keskiarvo eivät ole selvästi toisistaan riippuvaisia.

Yksiasteotantaa vastaava otoksesta laskettu varianssi saadaan aproksimaationa seuraavasti:

$$(2.9) \qquad V(\bar{x}'') = (1 - \frac{n}{N}) \frac{s^2_{b''}}{n} = (1 - \frac{n}{N}) \frac{1}{n} \frac{1}{n-1} \sum_i^n (\frac{M_i}{\bar{M}})^2 (\bar{x}_i - \bar{x}'')^2$$

$\bar{M} = \sum_{i=1}^{N} M_i / N$ voidaan korvata otoksesta lasketulla estimaatilla $\bar{M}' = \sum_{i=1}^{n} M_i / n$.

Suhde-estimaatin varianssi on pienempi kuin yksinkertaiseen ryväskeskiarvoon perustuvan, jos suhteen muodostavien termien välinen korrelaatio on suurempi kuin osamäärä $CV(M_i) / 2\ CV(M_i\bar{x}_i)$, jossa CV:t kuvaavat asianomaisten muuttujien variaatiokertoimia. Estimaattorilla (2.9) saadaan pienempi varianssi kuin harhattoman estimaatin varianssikaavalla, jos $M_i\bar{x}_i$:n ja M_i:n välinen korrelaatio on positiivinen ja suuruudeltaan > o.5. Yleensä kuitenkin estimaattori (2.9) antaa suurempia variansseja kuin vastaava yksinkertaiseen ryväskeskiarvoon perustuva, jos M_i:n ja termin $(\bar{x}_i - \bar{x}'')^2$ välinen korrelaatio on positiivinen.

Kaksiasteotannan mukaisessa kaavassa muodostuu rypäitten sisäisen varianssin lauseke pitkäksi ja laskennallisesti hankalaksi, mutta hyvänä aproksimaationa voidaan käyttää samantapaista ilmausta kuin estimaattorissa (2.7), jolloin kaava saa muodon

$$(2.10) \qquad V'(\bar{x}'') = (1 - \frac{n}{N}) \frac{s^2_{b''}}{n} + \frac{1}{nN} \sum_i^n (\frac{M_i}{\bar{M}'})^2 (1 - \frac{m_i}{M_i}) \frac{s_i^2}{m_i} .$$

Mikäli rypäistä on käytettävissä jonkin muun kuin mitattavan muuttujan osalta täydentävää tietoa, voidaan sitä käyttää hyväksi ottamalla avuksi varsinainen suhde-estimaatti. Jos otosrypäiden lukumäärä on suuri ja mitattavan muuttujan korrelaatio apumuuttujaan nähden korkea, voi suhde-estimaatti olla muita estimaatteja tehokkaampi. Tämä edellyttää usein kuitenkin rypäiden suuruuden huomattavaa vaihtelua.

2.3. Ryväsotannan tehokkuus

Ryväsotannan tehokkuus verrattuna muihin otantamenetelmiin riippuu, paitsi perusjoukon rakenteesta, myös rypäiden ja niiden elementtien lukumääristä ja keskinäisistä suhteista. Yleensä populaatioestimaatti on sitä tarkempi, mitä pienempi ryväs on. Tehokkuus vähenee melko nopeasti ryväskoon suuretessa.

Rypäitten rakenteeseen voidaan useimmiten vaikuttaa niiden konstruoinnin yhteydessä. Tässä mielessä ryväsotannalla on yhtymäkohtia ositetun otannan kanssa. Kuitenkin ne periaatteet, joiden avulla saadaan aikaan tehokas ositus, ovat huonoja ryväsotannassa käytettyinä. Mitä enemmän rypään elementit ovat toistensa kaltaisia, sitä paremmin rypäät soveltuvat ositetun otannan ositteiksi ja sitä huonommin käytettäviksi otantayksikköinä.

Käytännön tutkimuksissa, varsinkin jos rypäät konstruoidaan maantieteellisen sijainnin perusteella, ei yksityistä ryvästä voi pitää populaation elementeistä otettuna satunnaisnäytteenä. Tavallisesti samaan rypääseen kuuluvat elementit muistuttavat tutkittavan ominaisuutensa puolesta enemmän toisiaan kuin muihin rypäisiin kuuluvia elementtejä. Tästä johtuen rypäiden välinen hajonta pyrkii olemaan suurempi kuin niiden sisäinen hajonta. peruste? Siten ryväsotantaan perustuva varianssi on tavallisesti suurempi kuin elementtien lukuisuuden suhteen yhtä suuren yksiasteisen satunnaisotannan.

Seuraavassa tutkitaan ryväsotannan tehokkuutta niin sanotun intra-class-korrelaation avulla. Tämä korrelaatio, josta tullaan käyttämään nimeä sisäkorrelaatio, mittaa ryväsotannan kannalta keskeisen tekijän, rypäitten sisäisen homogeenisuuden (tai heterogeenisuuden), astetta. Vertailukohtana on osittamaton, yksiasteinen satunnaisotanta.

Koska aikaisemmin on oletettu rypäät kokonsa puolesta erisuuruisiksi, on tämä otettava huomioon myös sisäkorrelaatiota laskettaessa. Varianssianalyysin termein voidaan otoksesta laskettu korrelaatio ilmaista seuraavassa muodossa[1]:

---

1) Snedecor: Statistical Methods, Iowa 1962, s. 284.

$$(2.11) \qquad \rho = \frac{s_b^2 - s_w^2}{s_b^2 + (\bar{m} - 1)s_w^2} \quad , \text{ jossa}$$

s_b^2 = rypäiden välinen varianssi varianssianalyysissä,
s_w^2 = rypäiden sisäinen varianssi varianssianalyysissä,
$\bar{m}$ = otosrypäitten keskimääräinen elementtien luku.

Otettaessa termejä varianssianalyysistä on muistettava, että kyseessä on balansoimaton tapaus ja sovellettava sen mukaista menettelyä.[1] Kun m_i vaihtelee suuresti, saattaa otosrypäiden keskikoon laskeminen aritmeettisena keskiarvona aiheuttaa sisäkorrelaatioon virhettä. Sen vuoksi on kehitetty erikoinen menettely, jolla saadaan ryväskoon keskiarvo seuraavasti[2]:

$$\bar{m} = \frac{1}{n-1} \left(\sum_i^n m_i - \frac{\sum_i m_i^2}{\sum_i m_i} \right) .$$

Sisäkorrelaation arvo riippuu rypäitten välisen ja niiden sisäisen varianssin suhteesta. Jos $s_w^2 = 0$, toisin sanoen rypään kaikki elementit ovat mitattavaan muuttujaan nähden yhtä suuria, saa korrelaatio arvon yksi. Jos taas kaikkien rypäitten keskiarvot ovat yhtä suuret, saa korrelaatio miniminsä ja määräytyy yhtälöstä

$$\rho = \frac{-1}{\bar{m} - 1} \quad , \text{ jolloin}$$

sen arvo riippuu $\bar{m}$:n suuruudesta.

Sisäkorrelaation arvo on siis lähellä +1:tä, jos rypäiden välinen varianssi muodostaa huomattavan osan kokonaisvarianssista, ja lähellä nollaa oleva positiiviluku tai jopa negatiivinen, jos pääosa varianssista muodostuu rypäiden sisäisestä vaihtelusta. Jälkimmäisessä tapauksessa rypäät usein ovat tehokkaita otantayksiköitä.

---

1) Asiaa on käsitelty mm. teoksessa Seeger: Variance Analysis of Complete Designs, Uppsala 1966, s. 64-75.
2) Snedecor: mts. 269.

Mikäli rypäinä käytetään maantieteellisiä alueita, on sisäkorrelaatio tavallisesti positiivinen. Sen suuruus sen sijaan saattaa tällaisissa tapauksissa vaihdella missä tahansa nollan ja yhden välillä. Jos alkeisyksiköiden lukumäärä rypäässä on vähäinen ja ne ovat lähellä toisiaan, on odotettavissa voimakas positiivinen korrelaatio. Jos rypään koko taas on suuri ja elementit ovat alueellisesti hajallaan, voidaan päästä lähellä nollaa olevaan korrelaatioon.

Yksinkertaiseen satunnaisotantaan nähden voidaan ryväsotannan tehokkuus ilmaista tarkastelemalla niiden varianssien suhdetta. Seuraavassa kaavassa on ilmaistu ryväsotannan varianssi niin, että se on palautettavissa yksinkertaisen satunnaisotannan varianssiin.[1]

(2.12) $$V(\bar{x}) = (1 - \frac{n}{N}) \frac{V}{n\bar{m}} \left[1 + \rho(\bar{m} - 1)\right] ,$$ jossa ks. sivu 24

elementtejä kpl

$$V = s_b^2 + s_w^2 .$$ rypäitä kpl

$V(\bar{x})$ koostuu kolmen termin tulosta, joista viimeinen on ryväsotannan tehokkuuden kannalta mielenkiintoisin. Jos $\bar{m} = 1$ ja, kuten on laita yksiasteisessa satunnaisotannassa, m on aina yksi, saa myös viimeinen termi arvon yksi, ja $V(\bar{x})$ supistuu normaaliksi yksinkertaisen satunnaisotannan varianssiksi. Jos $\bar{m} > 1$, riippuu varianssin arvo olennaisesti sisäkorrelaatiosta. Täten voidaan termiä $\rho(\bar{m} - 1)$ käyttää mittaamaan sitä otosvarianssin muutosta, joka aiheutuu rypäiden käyttämisestä otantayksikköinä elementtien asemasta.

Vaikka ρ normaalisti pienenee $\bar{m}$:n kasvaessa, on sen muuttuminen suhteellisesti hitaampaa kuin $\bar{m}$:n. Näin ollen ryväskoon suureneminen johtaa positiivisen sisäkorrelaation vallitessa otosestimaatin varianssin kohoamiseen. Jos otoskoko alkeisyksikköjen suhteen on etukäteen kiintiöity, saa varianssi maksiminsa, kun rypään kaikki elementit otetaan näytteeseen, ja miniminsä, kun vain yksi elementti kustakin rypäästä sisällytetään otokseen. Tästä johtuu, että jos otannan kustannukset määräytyvät yksinomaan alkeisyksikköjen lukumäärän perusteella eikä esimerkiksi matkakustannuksilla ole merkitystä, päästään parhaaseen tulok-

1) Sukhatme: mts. 247.

seen, kun perusjoukkoa ei jaeta rypäisiin.

2.4. Ryväsotannan käyttö hakkuupoistuman määrittämisessä

Kaikki tarkastelut ulotetaan seuraavassa koskemaan koko maan laajuisia muuttujia. Käytännössä tarvitaan melkein aina tietoja myös osa-alueista, mutta siihen ongelmaan ei tulla puuttumaan muuten kuin hajanaisin esimerkein ja viittauksin.

2.4.1. Otantakehikko

Ensimmäiseksi on ratkaistava kysymys rypäiden konstruoinnista. Tällä seikalla on ~~hyvin paljon~~ keskeinen merkitys~~tä~~ parhaaseen mahdolliseen lopputulokseen pyrittäessä. Pääkriteereinä tulevat olemaan toisaalta tulosten tarkkuus ja toisaalta tutkimuksen kustannukset.

Metsälöiden rajaaminen rypäisiin ilman ristiriidatonta, ulkopuolista perustetta on sattumanvaraista. Metsätaloudellisissa tutkimuksissa tähän suuntaan meneviä kokeiluja on tehty muun muassa siten, että yksiasteisen satunnaisotannan perusteella saatuun näytemetsälöön on liitetty niin sanottu parimetsälö.[1] Parimetsälöllä tarkoitetaan kutakin varsinaista näytemetsälöä lähinnä sijaitsevaa metsälöä. Tällöin muodostuu otantayksiköksi kahden metsälön muodostama ryväs. Tämä menettely on perusteiltaan epävarma, sillä ryväs määräytyy siinä näytteen eikä populaation perusteella. Näin ollen eivät kaikki alkeisyksikköinä olevat metsälöt ole yksiselitteisesti luettavissa kuuluviksi vain yhteen rypääseen, mitä on pidettävä perusvaatimuksena alueotantaa tehtäessä.

Niinpä ongelmaa on lähestyttävä joltain muulta pohjalta. Erään mahdollisuuden tarjoaa karttojen käyttö. Voidaan ajatella koko maa jaetuksi kartan perusteella yhtä suuriin rypäisiin, jotka muodostavat otannan yksiköt. Tarkemmin ratkaistavaksi jää rypäiden muoto ja koko.

Ruotsissa suoritetuissa tutkimuksissa on alueotannan yhteydessä havaittu rypään muodon riippuvan osittain käytetystä poi-

---

1) Holopainen: Suomen metsien luovutusmäärä, Silva Fennica 97, Helsinki 1959, s. 8.

inventointikokeiluissako?

mintamenetelmästä.[1] Yleensä neliön muotoinen ryväs ei ole optimaalinen, mutta lähes kaikissa tapauksissa sen tehokkuus on hyvin lähellä optimaalimuotoisen rypään tehokkuutta. Jos ajatellaan kartan perusteella tehtävää rypäiden perusjoukon konstruoimista, on jo käytännöllisistä syistä suorakulmainen ryväs paras. Myös se, että ruotsalaisen Matérnin mukaan neliömuoto on parempi kuin muut suorakulmiot[2], on ollut painamassa päädyttäessä neliön muotoiseen rypääseen.

Itse konstruoiminen käy päinsä varsin helposti, sillä kartat voidaan jakaa koordinaattiviivoilla neliön moutoisiin ja neliökilometrin suuruisiin alueisiin. Ratkaistavaksi jää siten alueen koko. Tähän ongelmaan palataan kuitenkin uudelleen luvussa 2.4.3.

Täsmällisesti neliön muotoiset alueet eivät maastossa muodosta mitään luonnollisia kokonaisuuksia. Niiden paikallistaminen ja mittauksen suorittaminen tiukasti rajoja noudattaen muodostuu hankalaksi ja kalliiksi. Siksi onkin löydettävä jokin muu peruste.

Tarjolla on kaksi vaihtoehtoa. Ensinnäkin voidaan metsäpalstoittaisten karttojen perusteella rajata yksityinen ryväs niin, että siihen otetaan ne metsäpalstat, joiden pinta-alasta kuuluu asianomaiseen rypääseen suurempi osa kuin mihinkään muuhun rypääseen. Näin tulee jokainen palsta olemaan yhdessä ja vain yhdessä perusjoukon rypäässä.

Menetelmän suurimpana haittana on, että mittauksen perustuessa haastatteluun joudutaan metsäpalstan omistaja joskus etsimään muualta kuin tutkimusalueelta, koska eräissä osissa maata saman omistajan hallinnassa olevat metsäpalstat erilaisten maanjakotoimitusten yhteydessä ovat joutuneet hajalleen eri paikkoihin. Myös se, että saatavissa olevat metsäpalstarajoja osoittavat kartat eivät aina ole ajan tasalla, aiheuttaa lisätyötä.

Toisena vaihtoehtona on valita neliön muotoisen perusrypään mittausyksiköiksi kaikki sen rajojen sisäpuolelle osuneet talouskeskukset. Tällöin myös niistä kukin kuuluu yhteen ja vain yhteen perusjoukon rypääseen. Mittauksen kohteena ovat kaikki kyseisten talouskeskusten hallinnassa olevat metsälöt. Omistajan tai halti-

---

1) Matérn: Spatial Variation. Meddelanden från Statens skogsforskningsinstitut, Band 49, Nr 5, Stockholm 1960, s. 72-92.
2) Matérn: mts. 124.

jan paikallistaminen ei tuota vaikeuksia, koska hänet tavoittaa rypään alueelta. Tämän menetelmän pahimpana haittana voidaan pitää sitä, että rypään sattuessa tiheään asutulle seudulle tulee talouskeskusten lukumäärä rypäässä olemaan suuri.

Otantamenetelmän luonteesta johtuen ei omistussuhteita ja metsälörajoja tarvitse tutkia muiden kuin näyterypäitten osalta. Perusjoukon otantakehikon laatimiseen riittävät pelkästään tiedot valtakunnan rajojen koordinaattimerkinnöistä. Tähän tarkoitukseen sopivia karttoja on saatavissa koko maan osalta. Näyterypäitten tarkkaan rajaamiseen tarvitaan kuitenkin tarkempia karttoja. Saatavissa on kuntien veroluokituskarttoja, joiden mittakaavat ovat 1 : 20 000 ja 1 : 10 000, sekä kylittäin ja tiloittain laadittuja karttoja, joiden mittakaavat ovat 1 : 8 000 ja 1 : 4 000. Lisäksi on koko maan osalta valmistumassa tuoreisiin ilmakuviin perustuvia topografikarttoja, joista saadaan tietoja mittakaavoissa 1 : 50 000 ja 1 : 20 000.

2.4.2. Alkeisyksiköiden määrittely

Ryväsotannassa voidaan rypäitä nimittää ensiasteisiksi otantayksiköiksi tai vain otantayksiköiksi, koska otanta kohdistuu niihin. Mittaus kuitenkin suunnataan rypäässä tavallisesti johonkin pienempään yksikköön, jota nimitetään alkeisyksiköksi tai otannan elementiksi. Sen valinta riippuu suurelta osalta tutkimuksen tarkoituksesta, ja alkeisyksiköitä voi olla enemmänkin kuin vain yksi.

Tässä tutkimuksessa voidaan ryväs määritellä tarkemmin edellisessä luvussa selostetulla kahdella vaihtoehtoisella menetelmällä. Kummassakin tapauksessa on mahdollisuus olla laskennassa spesifioimatta mitään alkeisyksikköä ja tyytyä vain pelkkään ryväskohtaiseen agregaattiin. Tämän onnistuminen edellyttää, että ryväs mitataan kokonaan tai että kaikissa rypäissä voidaan pitää otantasuhde vakiona.

On hyvin todennäköistä, että tyytyminen vain ryväsagregaattiin antaa tarkkuuden kannalta huonoja tuloksia. Tämän vuoksi onkin syytä ottaa tarkempi yksikkö alkeisyksiköksi, varsinkin kun tällä toimenpiteellä sopivasti suoritettuna ei käytännöllisesti katsoen lainkaan lisätä kustannuksia. Vaikka hylätään talouskeskusten käyttö, on valittavana jälleen kaksi rationaaliselta tuntuvaa vaihtoehtoa.

Ensinnäkin voidaan alkeisyksiköksi valita metsäpalsta. Palstojen koko kuitenkin vaihtelee huomattavasti, ja koska hakkuumäärä on voimakkaasti korreloitunut metsäpalstan pinta-alan kanssa, hajonta muodostuu voimakkaaksi. Tulosten korotusten suhteen saattaa syntyä vaikeuksia, sillä ajan tasalla olevien lukumääräluetteloiden saanti on perusjoukon muutosten johdosta vaikeata. Suhde-estimaatin käyttö siten, että toisena satunnaismuuttujana on metsälön pinta-ala, parantaa tulosten tarkkuutta. Samoin on regressioestimaatin laita.

Toisena mahdollisuutena on valita alkeisyksiköksi metsähehtaari. Yksityiseltä hehtaarilta ei haastattelututkimuksessa voida mittauksia suorittaa, mutta rypäittäisiä tuloksia laskettaessa ollaan kuitenkin selvillä rypääseen sisältyvien metsäpalstojen hakkuumäärän ohella siitä metsäpinta-alasta, jolta nämä hakkuut ovat lähtöisin. Tätä menetelmää käytettäessä on otokseen sisällytettävä koko ryväs, sillä rypään sisällä tapahtuvaa hehtaarien välistä vaihtelua on mahdoton selvittää. Tällöin otannasta muodostuu tavallaan kahdessa vaiheessa ryväsotanta, sillä yksityinen metsäpalsta muodostaa metsähehtaareista koostuvan rypään ja metsäpalstat taas otantayksikkönä olevan aluerypään.

Vaikka metsäpalstoja ei pidettäisikään otannan peruselementteinä, voidaan niitä käyttää ositukseen. Suoritetuissa tutkimuksissa on tosin havaittu, etteivät metsähehtaaria kohti lasketut hakkuumäärät poikkea merkitsevästi toisistaan erilaisissa metsälösuuruusluokissa. Silti tietojen käyttäjien kannalta tällainen ositus on hyödyksi. Tällöin täytyy kukin ryväs erikseen hajottaa ositteisiin. Tulokset voidaan laskea joko yhdistämällä ensin samassa rypäässä ositteet yhteiseksi estimaatiksi tai laskemalla ositteittaiset arvot rypäittäin ja yhdistämällä ositteet vasta tämän jälkeen.

2.4.3. Rypäitten koko ja lukumäärä

Rypäiden todellinen koko tulee vaihtelemaan rypäästä toiseen, vaikka otantakehikon pohjana olevat neliöalueet ovatkin alunperin samansuuruiset. Perustaksi on kuitenkin otettava neliöalueen koon määrääminen. Koska rypäät peittävät koko pinta-alan eivätkä vain metsäpinta-alaa, tulee rypäiden todelliseksi keskimääräiseksi suuruudeksi metsäpinta-alan kokonaispinta-alasta lasketun prosenttimäärän suuruinen osa perusrypään pinta-alasta. Vuosina 1960-63 toimeenpannussa valtakunnan metsien inventoinnissa on veroluokituksen mukaisen metsämaan osuus valtakunnan koko pinta-alasta saatu 64.5 prosentiksi.

Rypäiden lukumäärä perusjoukossa riippuu niiden koosta. Taulukossa 1 on laskettu perusjoukon rypäiden lukumäärät, todelliset keskisuuruudet ja metsälöiden keskimääräinen luku rypäässä perusrypään sivun pituuden vaihdellessa välillä 0.5 - 5.0 kilometriä.

mistä aineisto tähän ?

Perusrypään sivun pituus km	Rypäiden määrä populaatiossa	Rypään keskikoko ha	Metsälöiden luku keskim.
0.5	1 348 130	16	0.3
1.0	337 030	65	1.1
1.5	149 790	145	2.5
2.0	84 260	258	4.5
2.5	53 930	403	7.0
3.0	37 450	581	10.1
4.0	21 060	1 032	18.0
5.0	13 480	1 613	28.2

Taulukko 1.

Seuraavissa laskelmissa tullaan käyttämään metsälöitä metsäpalstojen asemasta, koska viimeksi mainittujen määristä ja tunnusluvuista ei ole saatavissa tietoja. Näin muodostuva kokonaismäärä on metsäpalstoista saatavaa pienempi, sillä metsälö koostuu monessa tapauksessa useammasta kuin yhdestä palstasta.

Kun ryhdytään ratkaisemaan rypäitten kokoa ja niiden lukumäärää otoksessa, on ensin päätettävä, otetaanko mukaan kaikki yksityisen rypään elementit vai tehdäänkö osaotos. Koska rypäät mittauksellisesti koostuvat metsäpalstoista, ei rypään jakamista osiin voi tehdä muulla perusteella kuin metsäpalstarajoja noudattaen. Palstojen koot vaihtelevat huomattavasti, ja tämä tekee mahdottomaksi saada aikaan joka rypäässä edes likimain sama alaotoksen otantaosuus. Kun lisäksi ryväsotannan valitsemisen eräänä päävaikutteena on ollut kustannusten alentaminen käyttämällä hyväksi lähekkäin olevia mittauspisteitä, on otosrypäät otettu mukaan kokonaisina. Samalla saadaan eliminoiduksi rypään sisäinen varianssikomponentti.

Koska elementit, olivat ne sitten metsäpalstoja tai metsähehtaareja, ovat ryhmittyneet lähekkäin, voidaan rypäiden sanoa olevan kompakteja.[1] Kompaktin rypään optimikoon määrääminen on hankalampi tehtävä kuin kaksiasteisen otannan toisen asteen otosyksiköiden keskimääräisen luvun laskeminen.[2] Tämä johtuu siitä, että kompaktissa rypäässä sisäkorrelaatio vaihtelee rypään koon muuttuessa. Siksi on tarpeellista estimoida tämän korrelaation ja rypään keskikoon välinen riippuvuus.

Otoskoon optimaalinen määrääminen tähtää joko kustannusten minimointiin etukäteen annetun tarkkuusvaatimuksen puitteissa tai tarkkuuden maksimointiin annetuilla kustannuksilla. Otoksen koon ja samalla myös sen rakenteen määrääminen on siis luonteeltaan optimointitehtävä, jossa tarvittava informaatio perustuu toisaalta kustannusfunktioon ja toisaalta estimoitavien tunnuslukujen vaihteluun ja vaadittavaan tarkkuuteen. Kokonaiskustannuksille on usein asetettu määrätty tavoite. Samoin on tiettyjä etukäteisvaatimuksia tulosten tarkkuuteen nähden. Sen sijaan otannan eri vaiheiden keskinäisistä kustannuksista ja estimoitavien muuttujien vaihtelusta saatavat tiedot ovat tavallisesti epätarkkoja tai saattavat puuttua kokonaan.

Sopivan kokoinen esitutkimus antaa usein tarvittavaa informaatiota otoskoon määräämiseen. Esillä olevassa tapauksessa ei tällaista koetutkimusta ole ollut tilaisuutta suorittaa. Sen sijaan on käytettävissä tietoja muulla tavalla toteutetuista otantatutkimuksista, joissa estimoinnin kohteena on ollut osittain sama muuttuja kuin tässä.

Tuotettavan informaation tarkkuusvaatimus perustuu melkein aina harkintaan. Sen määräämisessä lähdetään tavallisesti liikkeelle koko tutkittavan alueen estimaatin tarkkuudesta. Esimerkiksi Holopaisen Suomen metsien luovutusmäärää koskevassa tutkimuksessa oli pyrkimyksenä päästä metsänhoitolautakunnittain alle 10 %:n suhteelliseen keskivirheeseen. Koko maan osalta muodostui

---

1) Kendall and Buckland: A Dictionary of Statistical Terms, London 1966, s. 46.

2) Hansen, Hurwitz and Madow: mts. 307.

tällöin keskivirheen prosenttiosuudeksi keskiarvosta 1.43 %.[1] Seuraavassa tullaan laskelmat perustamaan koko maan osalta kahden prosentin suhteelliseen keskivirheeseen hakkuumäärän osalta.

Aluksi määritellään kustannusfunktio. Yksinkertaisessa, lineaarisessa muodossa se voidaan ryväsotannassa esittää muuttuvien kustannusten osalta seuraavasti[2]:

$$C = c_1 n + c_2 n \bar{M} \text{, jossa} \tag{2.13}$$

C = muuttuvat kokonaiskustannukset,
c_1 = rypääseen liittyvät kustannukset,
c_2 = alkeisyksikköön liittyvät kustannukset.

Käytännössä c_2 edustaa luettelointiyksikköön, tässä tapauksessa metsäpalstaan, kohdistuvia kustannuksia. Tämä on varsin ymmärrettävää, sillä haastattelukustannukset esimerkiksi hehtaaria kohti vaihtelevat huomattavasti riippuen siitä, koostuvatko hehtaarit suuresta vai pienestä metsäpalstasta. Yksityisen palstan kohdalla kustannukset sen sijaan eivät suurestikaan riipu sen koosta, koska havainnointi oletetaan pääosiltaan suoritettavaksi haastattelun avulla.

Rypäästä toiseen siirtymisestä aiheutuviin kustannuksiin vaikuttavat monet tekijät. Olettamalla otosrypäät jakautuneiksi säännöllisesti tutkittavalla alueella, voidaan näiden tekijöiden vaikutusta kuvaamaan lisätä kustannusfunktioon termi $c_o \cdot n^{\frac{1}{2}}$, jossa c_o on empiiristä tietä määrättävissä oleva kerroin.[3] Mahalanobis on käytännön laskelmissa päätynyt tarkempaan arvoon, jossa $n^{\frac{1}{2}}$ korvataan termillä $n^{\frac{1}{2}} - n^{-\frac{1}{2}}$.[4]

Edellä on ryväsotannan tehokkuuden yhteydessä käsitelty sisäkorrelaation käyttökelpoisuutta otannan tehokkuuden selvittämisen apuna. Koska rypäät kokonsa puolesta vaihtelevat, vaihtelee jo tästäkin johtuen sisäkorrelaatio siten, että se pienenee ryväskoon suuretessa. Seuraavassa pyritään arvioimaan sisäkorre-

---

1) Holopainen: mts. 21.
2) Hansen, Hurwitz and Madow: mts. 272.
3) Hansen, Hurwitz and Madow: mts. 274.
4) Sukhatme: mts. 257.

laation suuruus erikokoisille rypäille. Tällöin perusyksikkönä pidetään jälleen metsälöä.

Suomessa tehdyistä empiirisistä tutkimuksista on käytettävissä vain niukasti informaatiota. E. Salon tutkimuksessa Etelä-Pohjanmaan ja Vaasan metsänhoitolautakuntien hakkuupoistumasta käytettiin rypäinä kuntia, jotka suuruutensa ja homogeenisuusasteensa puolesta eroavat tässä käytetyistä rypäistä.[1] Mittaus tapahtui siten, että yksityisellä metsälöllä laskettiin hakkuumäärä sen pinta-alaa kohden. Tämä toimenpide pienentää vaihtelua huomattavasti, koska pinta-alan ja hakkuumäärän välinen korrelaatio metsälöittäin laskettuna on voimakkaasti positiivinen. Etelä-Pohjanmaalla saatiin kahdesta eri alanäytteestä sisäkorrelaatiot, joiden arvot olivat 0.058 ja 0.031. Kumpaakin vastaavien rypäitten keskikoko oli 13 metsälöä. Vaasan metsänhoitolautakunnassa saatiin kahden alanäytteen sisäkorrelaatioiksi 0.055 ja 0.052, joita vastasivat 7 metsälön suuruiset rypäät.

Sukhatme on laskenut eräitä sisäkorrelaation arvoja Intiassa suoritetussa tutkimuksessa, jossa estimoitavana muuttujana oli vehnää viljellyn alueen suuruus. Yksityisessä kylässä olleista tutkimusalueista muodostettiin kompakteja rypäitä, joiden koko vaihteli. Rypään koon ja sisäkorrelaation välinen suhde muodostui seuraavaksi[2]:

Rypään koko	2	4	8	16
Sisäkorrelaatio	0.28	0.22	0.18	0.14

Myös Hansen, Hurwitz ja Madow ovat laskeneet sisäkorrelaatioita erilaisille muuttujille. Esimerkiksi kauppapuutarhojen suhteellista osuutta maanviljelystiloista kuvaavasta datasta saatiin seuraavia arvoja[3]:

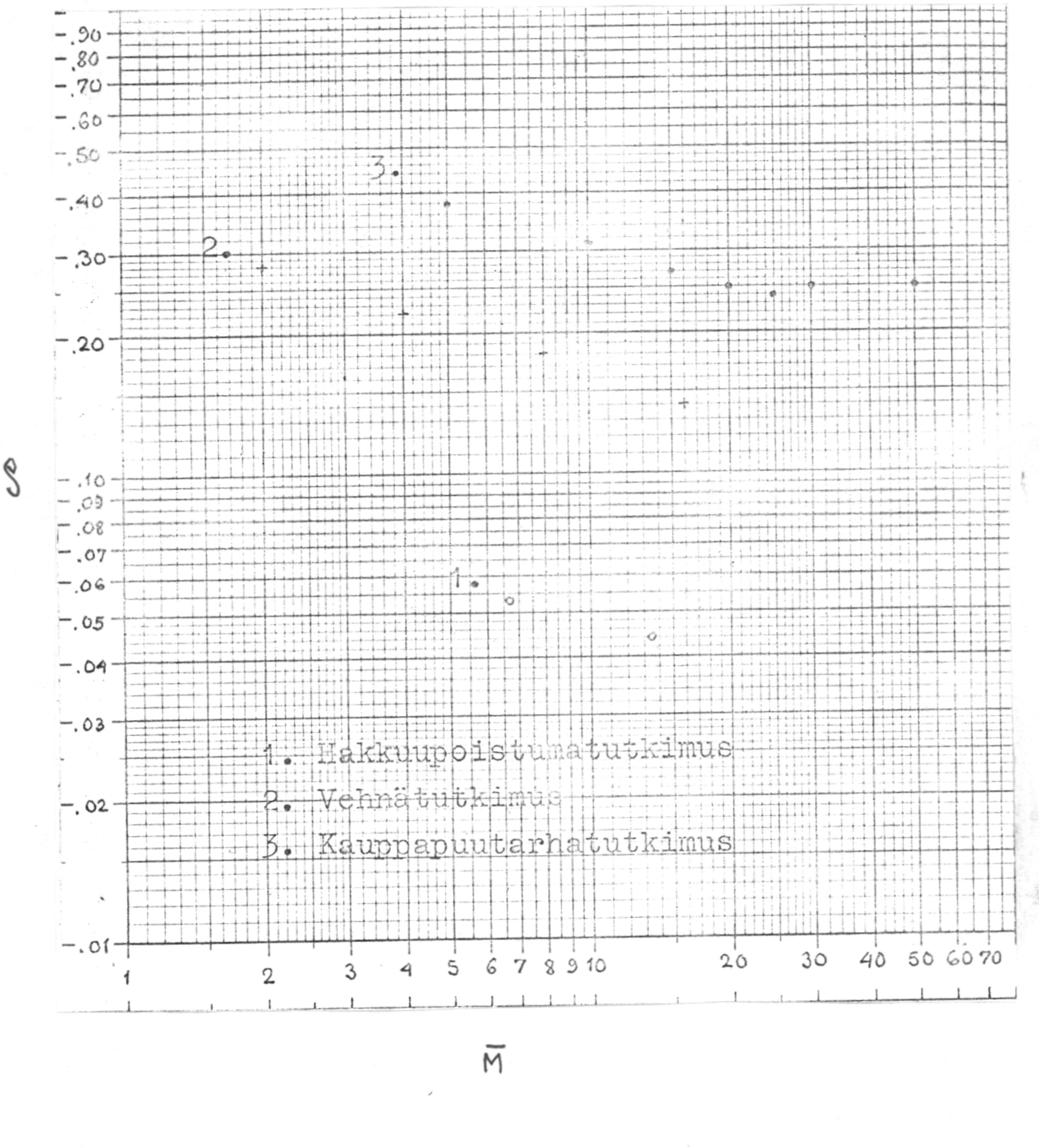
Rypään koko	5	10	15	20	25	30	50
Sisäkorrelaatio	0.38	0.31	0.27	0.25	0.24	0.25	0.25

1) Salo: Etelä-Pohjanmaan ja Vaasan metsänhoilautakuntien hakkuupoistuma hakkuuvuonna 1962-63. Julkaisematon tutkimus, Metsäntutkimuslaitos, Metsäekonomian tutkimusosasto.

2) Sukhatme: mts. 248,

3) Hansen,Hurwitz ja Madow: mts. 308.

Kun edellä esitetyt kolmesta eri lähteestä saadut luvut viedään logaritmipaperille, saadaan kuvion 1 mukainen esitys. Hakkuupoistumatutkimuksen kohdalla on eri näytteistä otettu aritmeettiset keskiarvot ja käytetty niitä.



Kuvio 1.

Sarjasta 2 muodostuu tämän logaritmimuunnoksen jälkeen erittäin lineaarinen kuvaus. Samaa on sanottava sarjan 3 alkuosasta. Sarjasta 1 on käytettävissä vain kaksi lukua, joten kuvaajan muodosta ei voi sanoa mitään. Jos riippuvuus kuitenkin oletetaan suoraviivaiseksi, näyttää tämän sarjan avulla saatavan kulmakertoimen arvo olevan lähellä kahden muun sarjan arvoja.

Jos merkitään regressiosuoran parametriarvoja a:lla ja b:llä, saadaan tavanomaiset yhtälöt lukusarjojen kuvaajille kaavasta

$$\log \varrho = a + b \log \bar{M},$$

josta saadaan ottamalla antilogaritmit

$$\varrho = a \bar{M}^{b} \tag{2.14}$$

Esitetyille kolmelle sarjalle tulevat parametrien arvoiksi seuraavat:

Sarja	a	b
1	0.086	- 0.262
2	0.301	- 0.327
3	0.478	- 0.191

Parametrin b arvot ovat melko yhdenmukaisia. Erilaisille muuttujille laskettu sisäkorrelaation pieneneminen ryväskoon kasvaessa on siten suuressa määrin samankaltaista. Harkinnanvaraisesti on seuraavissa laskelmissa valittu parametrin b arvoksi - 0.25. Hakkuumäärä yksityisen metsälön kohdalla on tällöin ajateltu laskettavaksi suhde-estimaattina metsälöön sisältyviä metsähehtaareja apumuuttujina käyttäen.

Seuraavana ongelmana on sisäkorrelaation tason määrääminen. Tätä tasoa kuvaa parametri a. Sen teoreettisesti mahdollinen vaihteluväli on -1:stä +1:een.

Ryväsotannan tehokkuutta käsiteltäessä todettiin, että käytettäessä rypäinä maantieteellisiä alueita sisäkorrelaatio normaalisti on positiivinen. Kun on kysymyksessä ryväs, jonka alkiot ovat alueellisesti hajallaan, on korrelaatiolla taipumus saada pieniä arvoja. Sen sijaan tarkasteltaessa kompakteja rypäitä, kuten Sukhatmen esittämässä esimerkissä, on odotettavissa voimakkaampi korrelaatio.

Koska etukäteen ei ole saatavissa tietoja tässä käsiteltävällä tavalla muodostetuista rypäistä, joudutaan jälleen turvautumaan harkintaan ja muihin tutkimuksiin. Ainoina tämän tutkimuksen kannalta relevantteina suomalaisina tietoina voidaan pitää aikaisemmin mainittuja Holopaisen luovutusmääriä ja Salon hakkuupoistumaa koskevia aineistoja. Sisäkorrelaatio on voitu laskea vain Salon aineistosta, ja sen arvo liikkuu todennäköisesti alemmalla tasolla kuin tässä kehikossa. Tämä johtuu siitä, että ensiksi mainitussa alkeisyksiköt sijaitsevat huomattavan hajallaan. Ilmeistä siis on, että Sukhatmen esimerkissä saatu parametrin a taso on parempi estimaatti. Niinpä a:lle onkin valittu arvo 0.30.

Näin saadaan ϱ:lle estimaatit erikokoisiin rypäisiin nähden, kun metsälöt ovat alkeisyksiköitä, lausekkeesta

$$\varrho = 0.3\ \bar{M}^{-0.25} .$$

Erisuuruisille perusrypäille saadaan nyt sivulla 17 olevan taulukon 1 mukaisesti seuraavat sisäkorrelaatiot:

Perusrypään sivun pituus	Metsälöiden luku keskim.	Sisäkorrelaatio	$\varrho(\bar{M} - 1)$
1.0	1.1	0.292	0.029
1.5	2.5	0.239	0.359
2.0	4.5	0.206	0.721
2.5	7.0	0.184	1.104
3.0	10.1	0.168	1.529
4.0	18.0	0.146	2.482
5.0	28.2	0.130	3.536

Taulukko 2.

Taulukon 1 vaihtoehdoista on ensimmäinen jätetty pois, sillä siinä metsälöiden lukumäärä ryvästä kohti on alle yhden. Koska rypäiden homogeenisuuden asteesta on nyt käytettävissä suuntaa antavaa tietoa, ~~joka luonteeltaan on kylläkin suuressa määrin hypoteettista,~~ voidaan otoksen optimirakenteen määräämisessä käyttää hyväksi sivulla 12 olevaa kaavaa (2.12). Jos oletetaan otantaosuus pieneksi, voidaan estimointi suorittaa lausekkeesta

$$(2.14) \qquad V(\bar{x}) = \frac{V}{n\bar{M}}\left[1 + \rho(\bar{M} - 1)\right] .$$

Vuoden 1966 ennakkotietojen mukaan oli hakkuupoistuman määrä 46.18 miljoonaa kiintokuutiometriä kuoretonta puuta. Metsähehtaaria kohti laskettuna luvuksi saadaan 2.124 k-m^3. Mikäli tarkkuusvaatimukseksi otetaan aikaisemmin esitetyn mukaisesti 2 %:n suhteellinen keskivirhe, saa $V(\bar{x})$ arvon 0.001805.

Termin V estimoimisessa käytetään seuraavassa hyväksi Salon hakkuupoistumatutkimuksen lukuja siitä huolimatta, että rypäinä siinä olivat kunnat. Keskimäärin oli kuntien metsäpinta-ala lähes 500 nelikilometriä, mikä merkitsee tässä esillä oleviin vaihtoehtoihin nähden noin 20-100 kertaa suurempien rypäiden käyttöä. Koska kuitenkin varianssit on laskettu hehtaaria kohti saaduista hakkuumääristä, ei rypään koolla ole niin suurta merkitystä kuin absoluuttisia hakkuumääriä käytettäessä.

Etelä-Pohjanmaan metsänhoitolautakunnan tuloksista saatiin V:n arvoksi 9.479. Koska sitä vastaava keskiarvo on 2.347, on tulosta pienennetty tämän luvun ja edellä esitetyn koko maata koskevan arvion 2.124 suhteiden neliöllä. Tällöin saatiin tulokseksi 7.763. Vaasan metsänhoitolautakunnassa tuli V:lle arvo 6.003, jota vastaa keskiarvo 2.015. Edellä esitetyn korjauksen jälkeen lopulliseksi tulokseksi saatiin 6.670. Vaihtelu näissä kahdessa lautakunnassa on Holopaisen luovutusmäärätutkimuksen mukaan suurempi kuin koko maassa keskimäärin. Variaatiokertoimeksi saadaan nimittäin mainitusta aineistosta Etelä-Pohjanmaalle 218 % ja Vaasan lautakunnalle 164 %, kun se koko maassa on 156 %.[1] Näin ollen on V:n estimaatiksi otettu Vaasan lautakunnan arvo alaspäin pyöristettynä, joten V saa seuraavassa arvon 6.0.

1) Holopainen: mts. 21.

aiheutuneeko Holopaisen otanta-järjestelyistä!?

Näin on saatu kaavaan (2.14) sisältyvien, laskennassa tarvittavien termien arvot. Kun yhtälö ratkaistaan n:n suhteen, saadaan erilaisia $\bar{M}$:n arvoja vastaavat n:n arvot seuraaviksi:

A metsän pinta	keskikoko $\bar{M}$ metsälöluku	ryppäiden luku n näytteessä	$n\bar{M}$	$100\frac{n}{N}$ %
1.0	1.1	3 110	3 421	0.9
1.5	2.5	1 807	4 518	1.2
2.0	4.5	1 271	5 720	1.5
2.5	7.0	999	6 993	1.8
3.0	10.1	832	8 403	2.2
4.0	18.0	643	11 574	3.1
5.0	28.2	535	15 087	4.0

Taulukko 3.

Tämän jälkeen voidaan kunkin $\bar{M}$:n ja sitä vastaavan n:n arvot sijoittaa kustannusfunktioon, joka rypäitten väliset matkakustannukset huomioon ottavana (ks. s. 19) saa muodon[1]

$$(2.15) \qquad C = c_o\sqrt{n} + c_1 n + c_2 n\bar{M} \ .$$

C:n arvon ratkaisemista varten tarvitaan vielä kertoimien c_i arvot. Oikeat absoluuttiset luvut eivät ole välttämättömiä, vaan eri kertoimien väliset suhteet riittävät hyvin, koska etukäteen ei ole asetettu vaatimuksia muuttuvien kokonaiskustannusten tasosta.

Kun käytettiin hyväksi aikaisempia aineistoja ja haastattelutottumusta omaavien henkilöiden kokemuksia[2], saatiin kertoimien arvoiksi

$$c_o = 235 \ , \quad c_1 = 3 \quad \text{ja} \quad c_2 = 6 \ ,$$

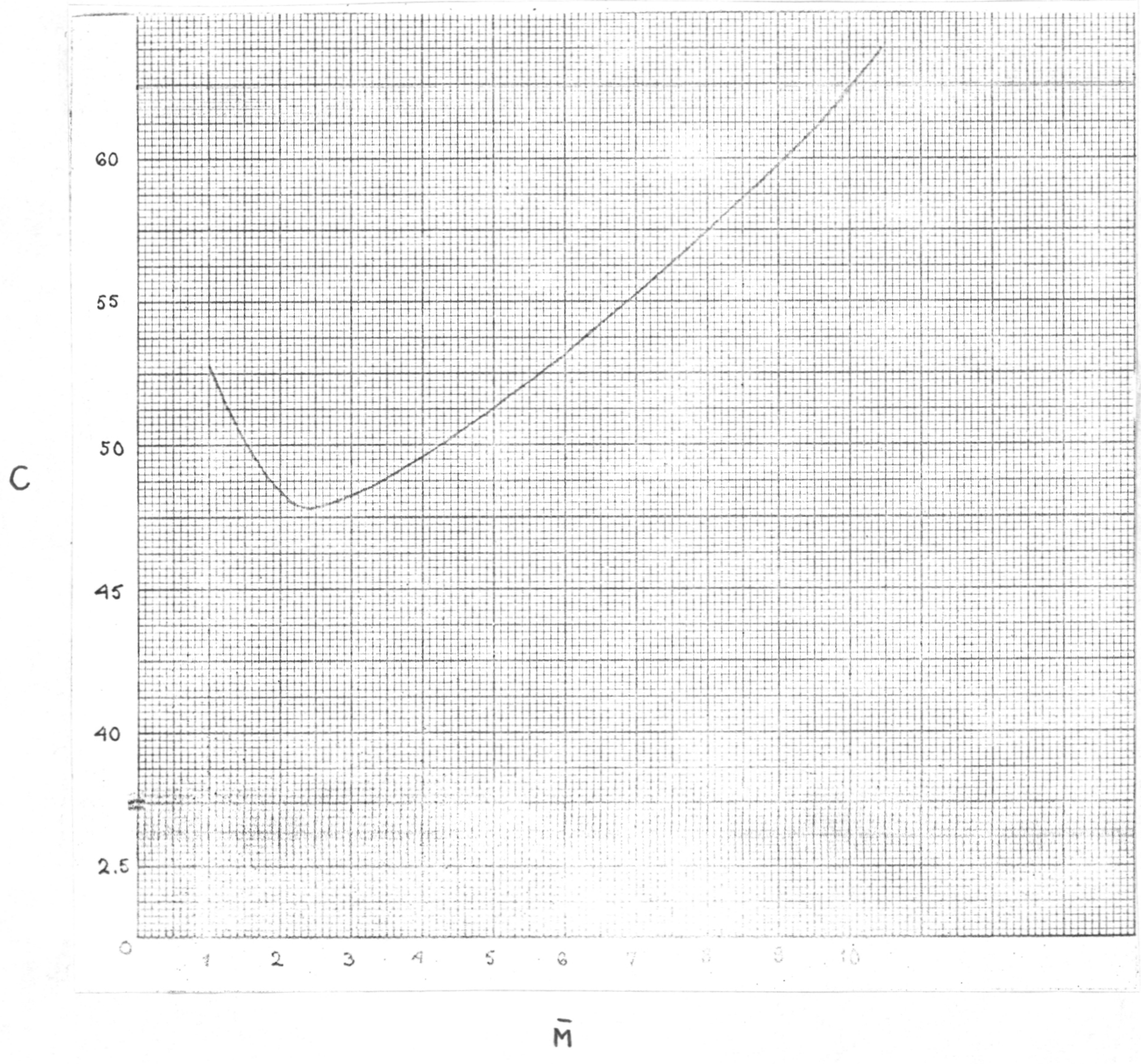
joten kustannusfunktio saa nyt muodon

1) Hansen, Hurwitz and Madow: mts. 310.

2) Pääosiltaan perustuvat saadut tiedot metsänhoitaja T. Huttusen antamaan aineistoon.

$$C = 235\sqrt{n} + 3n + 6n\bar{M} .$$

Kuvioon 2 on piirretty kuvaaja, joka ilmaisee kustannus-funktion perusteella saadun, samaa tarkkuustasoa vastaavien kustannusten suuruuden laskettuna erisuuruisille rypäille. Miniminsä funktio saavuttaa likimain $\bar{M}$:n arvolla 2.5, jota vastaa sivupituudeltaan 1.5 kilometriä oleva perusryväs. Kustannukset eivät nouse kovin jyrkästi ryväskoon suuretessa tästä. 2.0 kilometriä sivupituudeltaan olevaa perusryvästä vastaavat muuttuvat kustannukset ovat noin viisi prosenttia suuremmat. Tässä perusrypäässä keskimmärinen metsälöiden luku on 4.5.



Kuvio 2.

Yhden päivän aikana voi kokenut haastattelija suorittaa kompaktissa rypäässä tiedusteluja siten, että yhteen metsälöön käytetty aika siirtymisineen on noin 1 - $1\frac{1}{2}$ tuntia. Näin ollen 4 tai 5 metsälön läpikäymiseen menee aikaa noin 6 tuntia. Kun lisäksi siirtymiseen rypäästä toiseen menee aikaa keskimäärin $\frac{1}{2}$ tuntia ja rypään rajaamiseen ja metsälöiden alustavaan paikallistamiseen noin 1 tunti, muodostuu tällaiseen keskimääräiseen rypääseen käytetty aika yhdeksi työpäiväksi. Koska kustannukset tämän suuruisessa rypäässä poikkeavat optimista vain vähän, on päädytty perusrypääseen, jonka sivun pituus on 2.0 kilometriä.

Yhteenvetona esitetään seuraavassa valittua perusryvästä koskeva tärkein informaatio

Perusrypään sivun pituus	2.0	km
Rypäiden lukumäärä perusjoukossa	84 260	kpl
Rypään todellinen keskikoko	258	ha
Metsälöiden luku keskimäärin	4.5	kpl
Sisäkorrelaatio	0.206	
Rypäiden lukumäärä näytteessä	1 271	kpl
Metsälöiden lukumäärä näytteessä	5 720	kpl
Otantaosuus	1.5	%

3. SYSTEMAATTINEN OTANTA

3.1. Systemaattisen otannan periaatteista

Käytännön otantatyössä on tasavälisellä otannalla tärkeä asema. Ilmeistä on, että pääosa otantatutkimuksista tehdään käyttämällä hyväksi erilaisia systemaattisia menetelmiä. Tähän voidaan esittää kaksi pääsyytä. Ensinnäkin otantayksiköiden poimiminen ja paikallistaminen tutkimuskentällä on systemaattista otantaa käyttäen usein verrattomasti halvempaa ja helpompaa kuin satunnaisotannan periaatteita noudattaen. Lisäksi tietokoneiden yhä lisääntyvät käyttömahdollisuudet tarjoavat tasavälisen otannan yhteydessä etuja, joiden aikaansaama työn ja varojen säästö on tuntuva.

Toisena syynä esitetään usein näkemys, että systemaattisella otannalla saadaan näyte paremmin levitetyksi perusjoukkoon, ja näin otoksesta tulee edustavampi kuin satunnaisnäytteellä aikaansaadusta. Tämä näkökohta ei ole itsestään selvä eikä aina yleistettävissä. Tosin systemaattinen otanta monessa tapauksessa tuottaa tarkemman populaatioestimaatin kuin samankokoinen satunnaisnäyte, mutta otantaan sisältyvän satunnaisvirheen estimoiminen vaatii enemmän informaatiota kuin tavallisesti on saatavissa. Näin ollen voidaan harvoin olla täysin varmoja siitä, kuinka tarkka saatu estimaatti on.

Systemaattisen otannan kehittäjistä on mainittava erityisesti suomalainen J.W. Lindeberg. Hänen panokseensa metsätieteissä palataan luvussa 3.2.1. Amerikkalaiset W.G. ja L.H. Madow todistivat ensimmäisinä systemaattisen otannan ja satunnaisotannan keskiarvojen varianssien yhtäsuuruuden. Systemaattiselle otannalle ja otannalle koosta riippuvin todennäköisyyksin on L. Törnqvist kehittänyt teorian.[1]

Toistetun satunnaisalkuisen systemaattisen otoksen poimintakehikko voidaan kuvata porrasfunktiolla

$t =$
$n =$

(3.1) $$z_j(t,n) = z_{ij}\ , \quad \text{jos } nP_{i-1,j} < t \le nP_{ij}\ , \quad \text{jossa}$$

$$P_{ij} = \sum_{h=1}^{i} \frac{a_{hj}}{A}\ , \qquad \sum_{h} a_{hj} = A\ , \qquad z_{ij} = \frac{v_{ij}}{a_{ij}}\ .$$

?

Indeksi j kuvaa kehikon erästä permutoimalla saaduista järjestyksistä ja v_{ij} alkioiden havaintoarvoja. Suure a_{hj} / A on j:nnen kehikon h:nnen yksikön poimintatodennäköisyys.

Porrasfunktion t:lle annetaan satunnaisarvo τ, joka on jakautunut tasaisesti välille (0,1). Tämän jälkeen annetaan funktiossa $z_j(t,n)$ t:lle arvot τ, $\tau+1$, $\tau+2$, ..., $\tau+(n-1)$, joita vastaavat havainnot muodostavat otoksen. Näyte saadaan toistetuksi antamalla τ:lle erilaisia arvoja.

Systemaattisen otoksen harhaton keskiarvoestimaatti saadaan

1) Törnqvist: The Theory of Replicated Systematic Cluster Sampling With Random Start, Review of ICI, Vol. 31:1, 1963.

kaavasta

$$(3.2) \qquad z_j(\tau \mid n) = \frac{1}{n} \sum_{g=0}^{n-1} z_j(\tau + g, n)$$

ja summaestimaatti kaavasta

$$(3.3) \qquad v_j(\tau \mid n) = A\, z_j(\tau \mid n) .$$

Kun on estimoitava systemaattisen otoksen keskiarvon varianssi, nousee eteen se vaikeus, että käytettävissä on tavallaan vain yksi otospiste $z_j(\tau \mid n)$. Törnqvistin esittämä ratkaisu tähän on otoksen toistaminen, jolloin varianssille voidaan laskea harhaton estimaatti klassisella kaavalla1)

$$s_m^2(z_j(\tau \mid n)) = \frac{1}{m-1} \sum_{v=1}^{m} \left[z_j(\tau_v \mid n) - \bar{z}_j(T_m \mid n) \right]^2 , \text{ jossa}$$

$$\bar{z}_j(T_m \mid n) = \frac{1}{m} \sum_{v=1}^{m} z_j(\tau_v \mid n) \quad \text{ja}$$

~~os~~itetta
tpc

m = toistojen lukumäärä.

syst. otannan tehokkuuden
vertailu

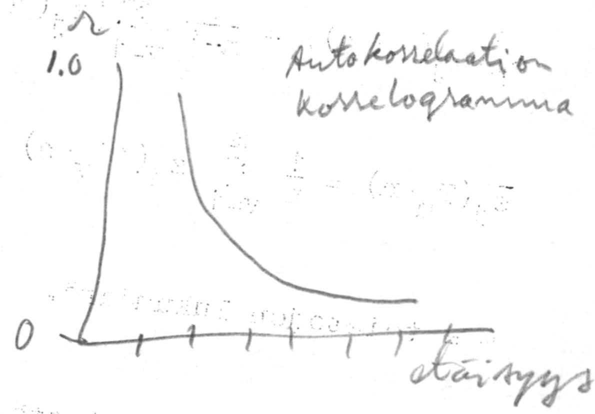
$$s^2 = \frac{1}{m} s_m^2(z_j(\tau \mid m))$$

Systemaattista otantaa käytettäessä on mahdollista vaikuttaa otannan tehokkuuteen seuraavilla toimintaparametreilla2):

a) otosrypäiden todennäköisyyspainojen valinta,
b) rypäiden konstruointi,
c) rypäiden järjestyksen valinta,
d) havaittavien rypäiden lukumäärän valinta,
e) näytteen toistojen lukumäärän valinta.

Poimintatodennäköisyydet on edullista valita niin, että ne ovat verrannollisia mitattavien suureiden arvoihin. Täten esimerkiksi hakkuupoistuman määrityksessä todennäköisyydet voidaan

1) Törnqvist: mts. 18.
2) Törnqvist: mts. 20.


r
1.0
0
Autokorrelaation
korrelogramma
etäisyys

valita rypään metsäpinta-alan eli todellisen koon mukaan. Jo aikaisemmin on todettu, että parhaaseen tulokseen päästään, kun rypäät ovat mahdollisimman heterogeenisia ja muodostavat siten pienoiskuvan perusjoukosta. Rypäiden järjestys on edullinen silloin, kun edellä esitetty porrasfunktio on mahdollisimman tasainen. Tällaiseen tavoitteeseen päästään, jos rypäät asetetaan metsäpinta-alan mukaiseen suuruusjärjestykseen. Havaittavien rypäiden ja näytteen toistojen lukumäärän kriteerinä on riittävän tarkkuuden saavuttaminen.

Ennen kuin tasavälistä otantaa ruvetaan soveltamaan, on selvitettävä perusjoukon rakenne ja luonne. Tärkeäksi muodostuu tällöin se periaate, jonka mukaan perusjoukon otantayksiköt on järjestetty. Pääsääntönä on, että järjestys ei ole määräävänä tekijänä otoksen koostumuksessa. Erityisesti on varottava, ettei pääse muodostumaan periodista vaihtelua esimerkiksi sinikäyrän muodossa. Jos tällöin otantaväli on heilahduksen tai sen monikerran pituinen, antaa systemaattinen otanta vääristyneen tuloksen.

Autokorreloituneissa perusjoukoissa systemaattinen otos on tehokkaampi kuin satunnaisotos, jos etäisyyden funktiona piirretty korrelogramma on ylöspäin kovera. Erityisesti alueellisissa tutkimuksissa autokorrelaatio saattaa pienillä etäisyyksillä olla hyvinkin voimakas. Tätä kysymystä on tutkinut ansiokkaasti mm. Matérn.[1]

Jos perusjoukon järjestys on satunnainen ja lähtöarvo valitaan satunnaisesti, voidaan helposti osoittaa, että systemaattisen otoksen otosyksiköiden aritmeettinen keskiarvo on perusjoukon keskiarvon harhaton estimaatti. Varianssin estimoiminen sen sijaan tuottaa hankaluuksia. Periaatteessa systemaattisen näytteen otosyksiköitä voidaan pitää yhtenä satunnaisesti valittuna otosrypäänä, jolloin otoksesta ei voida varianssiestimaattia muodostaa. Asia muuttuu toiseksi, jos käytetään toistoja, kuten Törnqvist on esittänyt.

Tavallisesti systemaattisen otoksen varianssi lasketaan ikäänkuin kyseessä olisi satunnaisotanta. Tällöin voidaan käyttää seuraavaa kaavaa[2]:

---

1) Matérn: mts. 51-68.

2) Sukhatme: mts. 431.

$$(3.4) \qquad V(\bar{x}) = (1 - \frac{1}{k}) \frac{s_s^2}{n}, \text{ jossa}$$

k on otantaväli ja n otosyksiköiden lukumäärä. Otosyksiköiden välistä varianssia kuvaavan termin s_s^2 odotusarvo voidaan lausua muodossa

$$s_s^2 = (1 - \frac{1}{N}) S^2 (1 - \rho) \quad , \text{ jossa}$$

S^2 on keskiarvon suhteen laskettu varianssi ja ρ on k:sta vaihtoehtoisesta rypäästä laskettu sisäkorrelaatio. Kaavan (3.4) perusteella saadaan tyydyttävä varianssiestimaatti vain, jos otannan yksiköt ovat perusjoukossa satunnaisessa järjestyksessä. Muussa tapauksessa riippuu estimaatin oikeellisuus sisäkorrelaatiosta. Jos se on suurempi kuin $-1/N-1$, antaa kaava keskimääräisesti aliestimaatteja todelliselle varianssille. Jos taas korrelaatio on pienempi kuin $-1/N-1$, saadaan yliestimaatteja.[1]

Myös toisenlaisille ideoille perustuvia estimaattoreita on kehitetty. Seuraava rakentuu kahden vierekkäisen havainnon erotukseen, ja se on harhainen kuten (3.4):kin.[1]

$$(3.5) \qquad V'(\bar{x}) = (1 - \frac{1}{k}) \frac{1}{2n} \frac{\sum_{i=1}^{n-1} (x_i - x_{i+1})^2}{n - 1} .$$

Kuten myöhemmin tullaan näkemään, perustuu Lindebergin metsänarvioimiseen kehittämä estimaattori analogiseen ideaan.

Mikäli kyseessä on suhde-estimaattori $q = x/y$, muuttuu kaava (3.5) muotoon[2]

$$(3.6) \qquad V(q) = (1 - \frac{1}{k}) \frac{1}{2n} \left[\frac{\sum_{i=1}^{n-1} (x_i - x_{i+1})^2}{(n - 1)\, \bar{x}^2} + \frac{\sum_i (y_i - y_{i+1})^2}{(n - 1)\, \bar{y}^2} - 2 \frac{\sum_i (x_i - x_{i+1})(y_i - y_{i+1})}{(n - 1)\, \bar{x}\, \bar{y}} \right]$$

---

1) Sukhatme: mts. 432.

2) Hansen, Hurwitz and Madow: mts. 506.

3.2. Systemaattisen otannan käyttö metsätieteissä

Laajimmin ja kauimmin on metsätieteissä tasavälistä otantaa käytetty metsänarvioimisen yhteydessä. Edustavista metsänarvioimismenetelmistä on ensimmäisenä esitetty linja-arviointi, jonka idea on se, että arviointi suoritetaan vetämällä metsään samanlevyisiä arviointilinjoja, jotka sijaitsevat yhtä suuren välimatkan päässä toisistaan ja kohtisuoraan maastokuvioiden yleistä esiintymissuuntaa vastaan.

Suomessa menetelmää käytettiin ensi kerran vuonna 1885[1], mutta varsinaisen käyttöönoton voidaan sanoa saaneen alkunsa vuodesta 1912, jolloin W. Cajanuksen ja Y. Ilvessalon toimesta suoritettiin Sahalahden ja Kuhmalahden pitäjien metsien arviointi. Ensimmäisen kerran kiinnitettiin tällöin huomiota autokorrelaatioon ja periodiseen vaihteluun, joista yhdessä käytettiin nimeä systemaattinen muuttuminen. Sen vaikutuksen huomioon ottamiseksi käytettiin seuraavaa menetelmää.[2]

Eri arvioimislinjoille saadut tutkittavan tunnuksen keskiarvot asetetaan pisteinä koordinaatistoon, jossa vaaka-akselilla on linjan numero. Jokaiselle linjalle saatu tulos katsotaan peruslinjasta lasketun etäisyyden funktioksi. Näin syntynyt pistesarja tasoitetaan graafisesti siten, että tasoitusviiva seuraa vain sellaisia vaihteluita, jotka näyttävät systemaattisilta. Keskiarvon varianssin laskemisessa määrätään kaikkien havaintopisteiden poikkeamat vastaavalla kohdalla olevasta tasoitusviivan pisteestä. Laskentakaavana käytetään seuraavaa:

$$V(\bar{x}) = \frac{1}{n} \sum_{i=1}^{n} \frac{l_i}{L}(x_i - x_i')^2 \quad , \qquad \sum_i l_i = L. \tag{3.7}$$

l_i = arviointilinjan i pituus,
x_i = havaittu tunnusluku linjalla i,
x_i' = tasoitusviivan arvo linjalla i,
n = linjojen lukumäärä tutkittavalla alueella.

---

1) Lappi-Seppälä: Linja-arvioimisesta ja sen tarkkuudesta. Metsätieteellisen koelaitoksen julkaisuja 7, Helsinki 1924, s. 13.
2) Ilvessalo: Tutkimuksia yksityismetsien tilasta Hämeen läänin keskiosissa. Acta Forestalia Fennica 26, Helsinki 1923, s. 24-30.

Tämän menetelmän heikkoutena on se, että tasoitusviivan määrääminen tapahtuu subjektiivisesti ja on siten vailla matemaattista pohjaa. Lisäksi keskiarvon varianssin laskentakaava on oikea vain siinä tapauksessa, että seuraavat ehdot on täytetty:

1) Luvut x_i ovat toisistaan riippumattomia satunnaisarvoja, joiden keskiarvo on x_i'.
2) Tutkittavan alueen todellinen tunnusluku on $\sum_i \frac{l_i}{L} x_i'$.
3) Arvojen x_i varianssit ovat kääntäen verrannollisia vastaaviin pituuksiin l_i.

Ehto 3) voidaan väistää, jos kaava (3.7) korvataan seuraavalla kaavalla, jossa on käytetty normaalia painotusmenetelmää:

$$V'(\bar{x}) = \sum_i \left(\frac{l_i}{L}\right)^2 (x_i - x_i')^2 . \tag{3.8}$$

Ruotsissa toimeenpantiin yllä selostettua tutkimusta edeltävänä vuonna Värmlannin läänin metsien linja-arviointi. Tulosten laskentaa varten linjat oli siinä yhdistetty kymmeneen ryhmään, ensimmäiseen linjat 1, 11, 21, 31, ..., toiseen 2, 12, 22, jne. Itse asiassa oli siis kyse toistojen aikaansaamisesta, vaikkakin se tapahtui jälkikäteen.

Keskiarvo ja sen varianssi laskettiin seuraavasti[1]:

$$\bar{x} = \sum_{j=1}^{10} \frac{L_j}{L} \bar{x}_j \quad , \quad \sum_j L_j = L . \tag{3.9}$$

$$V(\bar{x}) = \frac{1}{9} \sum_j \frac{L_j}{L} (\bar{x}_j - \bar{x})^2 . \tag{3.10}$$

L_j = linjaryhmän j linjojen yhteispituus,
$\bar{x}_j$ = linjaryhmän j havaintokeskiarvo.

---

1) Lindeberg: Über die Berechnung des Mittelfehlers des Resultates einer Liniertaxierung, Acta Forestalia Fennica 25, Helsinki 1924, s. 4.

Varianssikaavaan liittyy rajoittava ehto

$$\frac{L_1}{L}\sigma_1^2 = \frac{L_2}{L}\sigma_2^2 = \ldots = \frac{L_{10}}{L}\sigma_{10}^2 \quad , \text{ jossa}$$

σ_j^2 = linjaryhmän j varianssi.

3.2.1. J.W. Lindebergin linja-arvioinnin keskivirhe-estimaattori

Kun valtakunnan metsien ensimmäistä linja-arviointia ryhdyttiin vuonna 1921 suorittamaan, koettiin polttavaksi tarve saada kehitetyksi luotettavat menetelmät tulosten tarkkuuden estimoimiseksi. Tämän työn otti suorittaakseen J.W. Lindeberg.

Hän lähtee liikkeelle siitä perusideasta, että systemaattisen muuttumisen aiheuttamalta haitalta vältytään vertaamalla keskenään aina kahden vierekkäisen linjan tuloksia. Ideaansa Lindeberg ryhtyy kehittelemään kaavan (3.8) pohjalta.

Merkitsemällä $i \neq j$ saadaan

$$(3.11) \qquad E\left\{\frac{l_i^2}{L^2}(x_i - x_i')^2 + \frac{l_j^2}{L^2}(x_j - x_j')^2\right\} =$$

$$E\left\{\left[\frac{l_i}{L}(x_i - x_i') - \frac{l_j}{L}(x_j - x_j')\right]^2\right\}.$$

Kehittämällä tämän yhtälön vasemman puolen muotoon

$$E\left\{\left(\frac{l_i + l_j}{2L}\right)^2 (x_i - x_j)^2\right\} + \frac{l_i - l_j}{2L^2}(l_i\sigma_i^2 - l_j\sigma_j^2)$$

$$- \left(\frac{l_i + l_j}{2L}\right)^2 (x_i' - x_j')^2$$

ja erottamalla kaavasta (3.11) kaavan (3.8) mukaisen termin Lindeberg päätyy seuraavaan estimaattoriin[1]:

1) Lindeberg: mts. 14.

$$(3.12)\qquad E\left\{\sum_{i}^{n}\frac{l_i^2}{L^2}(x_i - x_i')^2\right\} =$$

$$E\left\{\frac{1}{2}\sum_{i}^{n-1}\left(\frac{l_i + l_{i+1}}{2L}\right)^2(x_i - x_{i+1})^2\right\} + R_1 + R_2 + R_3 .$$

$$R_1 = \frac{1}{2}\frac{l_1^2}{L^2}\sigma_1^2 + \frac{1}{2}\frac{l_n^2}{L^2}\sigma_n^2 .$$

$$R_2 = \frac{1}{2}\sum_{i}^{n-1}\frac{(l_i - l_{i+1})}{2L^2}(l_i\sigma_i^2 - l_{i+1}\sigma_{i+1}^2) .$$

$$R_3 = -\frac{1}{2}\sum_{i}^{n-1}\left(\frac{l_i + l_{i+1}}{2L}\right)^2(x_i' - x_{i+1}')^2 .$$

Jos pituudet l_i ja hajonnat σ_i eivät vaihtele kovin paljon, tulee R_1:n arvo olemaan noin n:s osa yhtälön (3.12) vasemmasta puolesta. Jotta R_2:n arvo nousisi merkittäväksi, täytyy tulojen $l_i\sigma_i^2$ vaihdella huomattavasti, ja vaihtelun tulee tapahtua vielä niin, että R_2:n summan osatermit eivät erimerkkisinä tasoita toisiaan. Tämä edellyttää kuitenkin niin erikoista tutkittavan alueen rakennetta, että R_2:n arvoa voidaan käytännössä pitää varsin pienenä.

Myös R_3:n arvo on yleensä pieni, sillä lukujen x_i' vaihteluiden pitää käytetyn tasoitusmenetelmän (ks. s. 32) mukaisesti vastata tunnuslukujen x_i systemaattisia muutoksia, ja koska tasoitusviiva on luonteeltaan jatkuva, on erotuksen $x_i' - x_{i+1}'$ arvo pieni.

Edellä olevaan nojautuen Lindeberg toteaa, että yhtälön (3.12) oikean puolen ensimmäiseen termiin verrattuna R-termien arvo on niin vähäinen, että ne voidaan jättää kaavasta pois. Näin päädytään kuuluisaan kaavaan, joka määrittelee linja-arvioinnin keskiarvon varianssin.

$$(3.13)\qquad V(\bar{x}) = \frac{1}{2}\sum_{i}^{n-1}\left(\frac{l_i + l_{i+1}}{2L}\right)^2(x_i - x_{i+1})^2 .$$

Tätä estimaattoria voidaan pitää pikemminkin liian suuria kuin liian pieniä arvoja antavana, sillä tarkassa kaavassa on R-termeistä negatiivisen R_3:n merkitys kaikkein suurin.

Linjojen lukumäärän ollessa vähäinen antaa kaava (3.13) systemaattisesti liian pieniä arvoja, koska termin R_1 merkitys korostuu. Tämän takia voidaan tekijä $\frac{1}{2}$ $(= \frac{n}{2n})$ korvata tekijällä $\frac{n}{2(n-1)}$, milloin linjojen lukumäärä < 20.

Vähäisen linjamäärän aiheuttaman epävarmuuden poistamiseksi Lindeberg esittää myös sellaisen menetelmän, että tutkittava alue jaetaan ositteisiin, joiden rajat leikkaavat käytettyjä arvioimislinjoja.[1] Tätä tapaa on käytetty mm. valtakunnan metsien inventoinneissa. Koko alueen varianssi saadaan siten tavanomaisesti ositteiden variansseista. Linjojen osittaminen tällä lailla johtaa yleensä parempaan tulokseen myös siitä syystä, että lukujen x_i satunnaiset vaihtelut tulevat tällöin suuremmiksi ja systemaattiset vaihtelut vaikuttavat näin ollen suhteellisesti vähemmän kuin ositusta käyttämättä.[2]

Edellä jo todettiin, että Lindebergin kaava johtaa pikemminkin liian suuriin kuin liian pieniin keskivirheisiin. Tätä kohtaan esitti norjalainen A. Langsaeter kritiikkiä ja toi julki, että olisi pyrittävä kaavoihin, jotka antavat pienemmän keskivirheen. Tätä varten hän kehitti oman kaavansa, joka perustuu yksityisille linjoille saatujen keskiarvojen toisiin erotuksiin.[3]

$$(3.14) \qquad V(\bar{x}) = \frac{n}{1.5\,(n-2)} \sum_{i=2}^{n-1} \left[\frac{(l_{i-1} + l_{i+1})\,l_i}{2\,L^2} \left(\frac{x_{i-1} + x_{i+1}}{2} - x_i \right)^2 \right]$$

Lindeberg myönsi tämän kaavan antavan omaa kaavaansa alhaisempia ja tarkempia arvoja, mutta totesi samalla, että toisten erotuksien laskenta käy paremmin päinsä, jos käytetään seuraavaa

1) Lindeberg: mts. 18.

2) Ilvessalo: Suomen metsät. Tulokset vuosina 1921-1924 suoritetusta valtakunnan metsien arvioimisesta. Metsätieteellisen koelaitoksen julkaisuja 11, Helsinki 1927, s. 295.

3) Langsaeter: Om beregning av middelfeilen ved regelmässige linjetaksering. Meddelelser fra det Norske skogsforsøksvesen, Oslo 1926, s. 32.

kaavaa[1]:

$$(3.15) \qquad V(\bar{x}) = \frac{n}{6(n-2)} \sum_{i=2}^{n-1} \left[\frac{l_i(x_{i-1} - 2x_i + x_{i+1})}{L} \right]^2 .$$

Jos estimaattoria halutaan tästäkin tarkentaa, voidaan käyttää kolmansiin tai neljänsiin erotuksiin perustuvia kaavoja, jotka tosin laskennallisesti ovat erittäin työläitä. Törnqvist on esittänyt lievän muunnoksen avulla laskentakaavat, jotka vastaavat toisia ja kolmansia erotuksia.[2]

3.3. Systemaattinen otanta metsälörypäisiin perustuvan näytteen konstruoinnissa

Hakkuupoistuman määritykseen edellä esitetty otantakehikko (ks. s. 13-14), jossa otosyksiköt alunperin koostuvat säännöllisistä neljän neliökilometrin alueista, soveltuu systemaattisessa otannassa käytettäväksi. Periodista vaihtelua tosin saattaa jonkin verran esiintyä siirryttäessä vuorotellen idästä länteen tai etelästä pohjoiseen. Periodin pituus ei kuitenkaan ilmeisesti ole säännöllinen.

Autokorrelaatiota esiintyy konstruoidussa kehikossa, kuten lähes kaikissa alueellisiin muuttujiin liittyvissä tapauksissa. Sen suuruudesta etäisyyden funktiona ei valitettavasti tässä vaiheessa ole kuvaa. Todennäköistä kuitenkin on, että korrelaatio on pienempi kuin jos kyseessä olisi kasvavan puuston määrän estimointi, sillä omistajat ovat etupäässä yksityisiä eikä keskitettyjä hakkuita ole vielä saatu aikaan suuressa mittakaavassa. Koska teollisuuden käyttämä raakapuu muodostaa kaksi kolmattaosaa koko hakkuupoistumasta, vaikuttaa tämä osa huomattavasti autokorrelaation suuruuteen, varsinkin kun loppuosa koostuu pääasiassa kiinteistöjen polttopuusta, jonka vaihtelu ei ole huomattava.

Teollisuudelle tarjottavan raakapuun määrään vaikuttaa suurelta osin puusta saatava kantohinta, joka puolestaan kokemuksen mukaan riippuu paljon metsälön sijainnista. Kuljetuskustannusten

---

1) Lindeberg: Zur Theorie der Liniertaxierung. Acta Forestalia Fennica 31, Helsinki 1927, s. 8.

2) Törnqvist : mts. 19.

vaikutusta kantorahaan on Suomessa selvitellyt L. Heikinheimo.[1] Ilmeisesti teollisuuskeskuksen sijainti on eräänä määräävänä tekijänä autokorrelaation muodostumiseen ja muuttumiseen etäisyyden funktiona.

Törnqvistin menetelmä, jossa on yhdistetty systemaattinen otanta ja otanta koosta riippuvin todennäköisyyksin, ei sovellu esillä olevaan tapaukseen, vaikka sen edut muuten ovatkin ilmeiset. Nimensä mukaisesti tekniikka vaatii tietoja perusjoukon rypäiden koosta. Tämä vaatimus ei mahdu tässä selostetun suunnitelman perusideoitten puitteisiin, sillä eräs keino, jolla tähdätään kustannusten alentamiseen, on , että perusjoukkoa ei tarvitse kokonaisuudessaan luetteloida. Sen sijaan toistojen suorittaminen käy periaatteessa päinsä, joskin tällöin rypäiden lukumäärän käyttö tehokkaana toimintaparametrina edellyttää todennäköisyyspainojen käyttämistä.[2]

3.3.1. Kaksiulotteinen systemaattinen otanta

Periaatteessa esillä oleva otantakehikko on kaksiulotteinen, kuten yleensä on laita maantieteelliselle perustalle rakennetuissa kehikoissa. Usein tällaisia kehikoita kuitenkin käsitellään siten, että yksiköt numeroidaan yksiulotteisen otannan suoritusta silmälläpitäen. Tyypillinen esimerkki on metsien linja-arviointi.

Kaksiulotteista otantakehikkoa voidaan tietysti käsitellä myös siten, että siitä otetaan kaksiulotteinen otos. Systemaattista otantaa käyttäen tämä käy hyvin päinsä. Seuraavassa tarkastellaan kahta vaihtoehtoa kaksiulotteisen systemaattisen otoksen suorittamiseksi neliön muotoisista alueista.[3]

Ensimmäinen menettely on seuraava: Valitaan satunnaislukupari, joka määrää ensimmäisen neliöalueen sisällä otannan kohteeksi tulevan pisteen. Tämä piste määrää neliöalueessa tietyt koordinaatit. Näiden koordinaattien perusteella valitaan sitten kaikista muista nelialueista otospisteet vastaavilta kohdilta.

---

1) Heikinheimo: Kantorahamalli. Metsätaloudellinen aikakauslehti n:o 12, Helsinki 1966, s. 1-3.

2) Törnqvist: mts. 20.

3) Cochran: Sampling Techniques, New York 1963, s. 228-29.

Toisessa vaihtoehdossa valitaan lisäksi (n-1)+(m-1) uutta satunnaislukua, missä n on horisontaalitasossa ja m vertikaalitasossa olevien neliöalueiden lukumäärä. n-1 lukua määrittelevät etäisyydet vaakatasosta ja m-1 lukua etäisyydet pystytasosta. Menetelmä on itse asiassa samanlainen kuin kokeiden suunnittelussa käytetty latinalaisen neliön periaate.

Käytännön tutkimuksissa on jälkimmäinen tapa havaittu yleensä paremmaksi. Tämän tutkimuksen kohteena olevaa otantakehikkoa ajatellen on kuitenkin päädytty edelliseen otantapisteiden määräämisen yksinkertaisuuden vuoksi. Lisäksi myöhemmin otannan toistamisen yhteydessä tulee esille seikkoja, jotka tukevat tätä valintaa.

Luvun 2.4.3. lopussa olevaa yhdistelmää hyväksi käyttäen päädytään seuraavaan kehikkoon. Perusideana on, että 2x2 kilometrin suuruisista perusrypäistä muodostetaan yhtä suuria neliön muotoisia alueita, joista kustakin otetaan näytteeseen yksi perusryväs. Neliöalueen suuruus määräytyy otokseen kiintiöityjen perusrypäiden määrän mukaan seuraavasti:

$$A^2 = \frac{84\ 260}{1\ 271} (\ 2\ \text{km}\ )^2 = 265\ \text{km}^2\ .$$

Tällöin saadaan neliöalueen sivun pituudeksi $\sqrt{265\ \text{km}^2}$ = 16.3 km, joten lopulliseksi pituudeksi tulee 16 km ja alueen suuruudeksi 256 km^2. Neljän neliökilometrin suuruisia perusrypäitä tähän alueeseen sisältyy 64 kpl, ja otosyksiköiden kokonaismäärä kohoaa tästä johtuen 1 317:ksi (vrt. s. 27). Samoin nousee metsälöiden odotettu luku näytteessä 5 936:een ja otantaosuus 1.6 %:iin.

Näyterypäiden valinta tapahtuu siten, että poimitaan aluksi kaksi satunnaislukua väliltä (1,8). Näistä toinen määrää otosrypään järjestyksen kussakin neliöalueessa horisontaalitasossa ja toinen vertikaalitasossa. Valinta voidaan periaatteessa tehdä ilman karttoja, kunhan on tiedossa valtakunnan rajalle sattuneiden neliöalueiden tarkat koordinaatit.

4. SYSTEMAATTISEN RYVÄSOTANNAN TUNNUSLUKUJEN LASKEMINEN HAKKUUPOISTUMAN MÄÄRITYKSEN YHTEYDESSÄ

Yhteenvetona edellä esitetyistä osamenetelmistä kootaan seuraavaksi hakkuupoistuman määrityksen yhteydessä tarvittavat tunnuslukuestimaattorit. Tämä tapahtuu siten, että ensin suoritetaan ongelman kuvaus ylhäältä alaspäin ja sen jälkeen esitetään estimaatit lähtien liikkeelle perustasosta.

Ylätason menetelmänä on systemaattinen otanta. Otoksen poimiminen tapahtuu kaksiulotteista maantieteellistä otantakehikkoa käyttäen ilman samana ajankohtana suoritettavia toistoja. Otokseen tulleista neljän neliökilometrin perusrypäistä otetaan lopulliseen näytteeseen ne metsäpalstat, jotka ovat kokonaan neliörajojen sisäpuolella, sekä ne palstat, joiden pinta-alasta otosrypääseen kuuluu suurempi osa kuin mihinkään muuhun perusjoukon rypääseen. Näin saaduista otosrypäistä suoritetaan mittaukset kaikista niihin sisällytetyistä metsäpalstoista. Saadut tunnusluvut lasketaan suhde-estimaatteina metsäpinta-alaan nähden.

Ennen kuin lopulliset laskentakaavat esitetään, on syytä puuttua valtakunnan rajojen ja metsättömien alueiden kuten vesien, peltojen, avosoiden ja asutuskeskusten muodostamaan mittausongelmaan. Valtakunnan rajoilla on meneteltävä siten, että perusjoukkoon kuuluvaksi luetaan sellainen perusryväs, jonka pinta-alasta vähintään puolet eli kaksi nelikilometriä kuuluu maan rajojen sisäpuolelle. Näin ollen rajalla sijaitsevat perusrypäät eivät ole kooltaan alkuperäisen suuruisia. Vastaavaa menettelyä on sovellettava käytettäessä maan sisällä osa-alueita kuten metsänhoitolautakuntia ja talousalueita.

Toisen ongelman muodostavat sellaiset rypäät, joissa ei ole metsää lainkaan. Käytettäessä metsäpinta-alalla painotettua otantajärjestelmää, ei tätä vaikeutta esiinny, sillä metsättömän rypään poimintatodennäköisyys on nolla. Suuria avosoita ja avomerialueita lukuunottamatta metsättömiä rypäitä ei kuitenkaan tule paljon esiintymään, sillä perusrypään koko, 400 hehtaaria, on niin suuri, että sellaisia avoalueita ei montaa tapaa edes järvialueilta, joilla saaret yleensä ovat metsäisiä. Koska kyseessä on pääosiltaan hakkuupoistuman estimoiminen, ja koska hakkuu kohdistuu metsään, ei analogisesti näyterypäitten konstruoinnin kanssa tällaisia rypäitä tulla ottamaan laskelmissa huomioon, vaan jos sellainen sattuu näytteeseen, se hylätään.

Lopuksi esitetään tulosten laskemista varten tarvittavat tunnuslukujen estimaattorit. Koska on päädytty suhde-estimaattiin, saadaan tutkittavan muuttujan arvo rypäässä yksinkertaisesti kaavasta

$$(4.1) \qquad q = \frac{\sum_{j=1}^{m_i} x_{ij} \,/\, m_i}{\sum_j y_{ij} \,/\, m_i} = \frac{\bar{x}_i}{\bar{y}_i} = \frac{x_i}{y_i}\,, \text{ jossa}$$

x = mitattava muuttuja (hakkuumäärä k-m^3),
y = metsähehtaari,
i = rypään tunnus,
j = metsäpalstan tunnus,
m = metsäpalstojen lukumäärä.

Rypäiden sisäinen varianssikomponentti eliminoituu pois, koska koko ryväs on mittauksen kohteena.

Tutkittavan alueen estimaattia laskettaessa ovat periaatteessa tarjolla ne kolme vaihtoehtoa, jotka on esitetty luvussa 2.2. Ensimmäinen perustuu ryväskohtaisten suhde-estimaattien tavalliseen aritmeettiseen keskiarvoon. Sille saadaan varianssi systemaattisen otannan tapauksessa joko kaavaan (2.3) perustuvasta kaavasta (3.4) tai Lindebergin idean mukaisesta kaavasta (3.5).

Toinen vaihtoehtoinen, kaavaan (2.5) perustuva estimaattori antaa harhattomia tuloksia, mutta sen varianssi muodostuu niin suureksi, ettei sitä sen takia käytännön laskelmissa kannata ottaa huomioon. Lisäksi sen käyttö vaatii perusjoukon osalta tietoja, joiden hankkiminen ja ajan tasalla pitäminen on työlästä.

Kolmas mahdollisuus perustuu kaavaan (2.8), joka on luonteeltaan suhde-estimaattori. Mikäli painoina M_i käytetään kappalemäärien asemasta metsähehtaareja, päädytään itse asiassa varsinaiseen suhde-estimaattoriin, sillä

$$(4.2) \qquad \bar{x} = \frac{\sum_{i=1}^{n} y_i q_i}{\sum_i y_i} = \frac{\sum_i^n x_i}{\sum_i y_i} = \frac{\bar{x}}{\bar{y}} = q\,.$$

Suhde-estimaattorille q saadaan systemaattisen otannan tapauksessa varianssi esimerkiksi kaavan (3.6) mukaisesti. Se antaa kuitenkin yliarvioita, jos rypäiden välinen autokorrelaatio on korkea.

Tämän kaltaisissa suhde-estimaateissa eivät tyhjät rypäät muodosta samankaltaista ongelmaa kuin muissa estimaateissa, sillä minkäänlaisia 0/0-tapaisia ryväsestimaatteja ei tarvitse käsitellä.

Lopullisesti päädytään tässä suosittamaan estimaattoreiden (4.2) ja (3.6) käyttöä niiden antamasta harhaisesta tuloksesta huolimatta, sillä rypäiden tavalliseen aritmeettiseen keskiarvoon perustuva estimaatti voi olla vielä harhaisempi, koska rypäiden koko vaihtelee huomattavasti.

Mikäli osa-alueille halutaan laskea estimaatteja, käy se päinsä esitettyjen kaavojen perusteella. Jos osa-alueet muodostavat keskinäisen vaihtelunsa perusteella sopivia ositteita, voidaan tulosten tarkkuutta parantaa laskemalla ensin tulokset niille ja yhdistämällä ne sitten koskemaan koko maan estimaatteja. Jos merkitään osa-alueen indeksiä l:llä, saadaan niistä estimaatit koko maalle normaaliin tapaan seuraavasti:

$$(4.3) \qquad q = \frac{\sum_{l=1}^{p} Y_l q_l}{\sum_{l} Y_l} \qquad \text{ja}$$

$$(4.4) \qquad V(q) = \frac{\sum_{l}^{p} Y_l^2 V(q_l)}{(\sum_{l} Y_l)^2} \quad , \text{ joissa}$$

p = osa-alueiden lukumäärä,
Y = osa-alueen metsähehtaarien kokonaismäärä.

Korotettaessa tuloksia vastaamaan perusjoukon kokonaisestimaatteja on menettely normaali. Suhde-estimaatti q kerrotaan perusjoukon metsähehtaarien kokonaismäärällä Y ja varianssi V(q) kerrotaan Y^2:lla. Osa-alueista korotetut tulokset saadaan yksinkertaisesti kaavojen (4.3) ja (4.4) osoittajista.

28.5 3.0

31.5

1968

1969

5. REGRESSIOMENETELMÄ

5.1. Regressioestimaattoreiden ominaisuuksista

Osituksen tapaan regressioestimaattorit on kehitetty lisäämään otannan tehokkuutta käyttämällä hyväksi tutkittavasta perusjoukosta saatavaa, estimoitavaan muuttujaan nähden rinnakkaista informaatiota. Yleisessä lineaarisessa muodossa voidaan otoksesta laskettu regressioestimaatti ja sen varianssi esittää seuraavasti[1]:

$$(5.1) \qquad \bar{x}_r = \bar{x} + b\,(\bar{Y} - \bar{y}) \,,$$

$$V(\bar{x}_r) = \left(1 - \frac{n}{N}\right) \frac{\sum_{i=1}^{N} \left[(x_i - \bar{X}) - b\,(y_i - \bar{Y})\right]^2}{n\,(N - 1)} \,, \text{ jossa}$$

varianssi saa miniminsä, kun b saa arvon

$$b = \frac{\sum_{i}^{N} (x_i - \bar{X})(y_i - \bar{Y})}{\sum_{i}^{N} (y_i - \bar{Y})^2} \,.$$

Tämän mallin käyttö edellyttää, että tietyt ehdot on täytetty. Ensinnäkin täytyy olla tiedossa apumuuttujan y todellinen populaatiokeskiarvo $\bar{Y}$. Toinen ehto on se, että muuttujien x ja y välinen riippuvuus on suoraviivainen, ja kolmas, että x:n varianssi keskiarvoonsa nähden on sama kaikilla y:n tasoilla.

Periaatteessa on mahdollista käyttää regressioestimaattoria yllä esitetyssä muodossa estimoimaan hakkuupoistumaa. Apumuuttujana y on tällöin metsähehtaarien lukumäärä, josta on valtakunnan metsien inventoinnin perusteella saatavissa sekä koko maata että osa-alueita koskevia tietoja.

Yksinkertaiseen satunnaisotantaan verrattuna regressioestimaatin varianssi on pienempi tai yhtäsuuri riippuen muuttujien

---

1) Cochran : mts. 191.

x ja y välisestä korrelaatiosta. Näiden kahden estimaatin varianssien suhde on likimain $(1 - r^2)$[1], joten tehokkuus laskee melko nopeasti korrelaation pienetessä. Suhde-estimaattoriin nähden antaa regressioestimaattori pienemmän varianssin lukuunottamatta tapausta, jossa y:n ja x:n välinen korrelaatio on yhtä suuri kuin niiden variaatiokertoimien suhde. Tällöin molemmilla estimaateilla on sama varianssi.[2]

5.2. Regressioestimaatin käyttö kaksoisotannassa

Kun apumuuttujan populaatiokeskiarvoa ei ole saatavissa, ei tavanomaiseen regressiomenetelmään perustuva otanta sovellu käytettäväksi. Tämän vuoksi on kehitetty ns. kaksoisotanta (double sampling), jossa aluksi otetaan suuri otos apumuuttujan y estimoimiseksi. Tästä otoksesta mitataan sitten tietyn suuruinen alanäyte, jossa mittaus kohdistetaan myös muuttujaan x, jolloin saadaan selville näiden kahden muuttujan välinen riippuvuus.

Erityisen hyvin tämäntapainen menettely soveltuu useana eri ajanjaksona toteutettavaan tutkimukseen, jolloin ajanjaksona t otetaan asianomaisesta muuttujasta normaali otos ja sen lisäksi samasta muuttujasta alaotos. Ajankohtana t+1 otetaan vain kyseinen alaotos, jossa tapahtuneen muutoksen perusteella estimoidaan kokonaisotoksessa tapahtunut muutos käyttäen ajankohdan t kokonaisotosta koskevia tietoja apumuuttuj na y. Tällaisessa useiden ajanjaksojen tutkimisessa halutaan tavallisesti tietoja seuraavista seikoista[3]:

1) $\bar{x}$:n muutos ajanjaksosta toiseen,
2) $\bar{x}$:n keskimääräinen arvo pitkänä aikavälinä,
3) $\bar{x}$:n tasoestimaatti tuoreimpana ajankohtana.

Yllä kuvatulla menettelyllä saadaan tasoestimaatti kaavan (5.1) perusteella. Mikäli ajankohdan t alkuperäisotos ja osaotos on valittu satunnaisotantaa käyttäen ja mitattavat elementit

1) Hansen, Hurwitz and Madow: mts. 459.
2) Hansen, Hurwitz and Madow: mts. 460.
3) Cochran: mts. 341.

ovat normaalisesti jakautuneet, saadaan keskimääräinen varianssi kaikille mahdollisille otoksille seuraavasti[1]:

$$(5.3) \qquad V(\bar{x}_d) = \left[1 + (1 - \frac{n'}{n}) \frac{1}{n' - 3}\right] (1 - r^2) \frac{s_x^2}{n'} + r^2 \frac{s_x^2}{n} .$$

Yhtälössä on n alkuperäisen otoksen ja n' alaotoksen otosyksiköiden määrä sekä s_x^2 estimoitavan muuttujan varianssi.

Jos muuttujat y_i eivät ole normaalisesti jakautuneet, muuttuu vain termin $1/(n' - 3)$ arvo. Tämän termin tarkan arvon suhteen esittää Cochran epäilyksiä. Mikäli n' on suuri, tulevat termit $(1 - n'/n)$ ja $1/(n' - 3)$ pieniksi, jolloin hakasulkujen sisällä oleva lauseke lähenee ykköstä, ja näin ollen ei termin $1/(n' - 3)$ arvolla ole käytännöllistä merkitystä.

Estimoitaessa muutos täysin samasta näytteestä tulee muutosestimaatin virhevarianssiksi tavanomainen

$$(5.4) \qquad V(\bar{x}_{t+1} - \bar{x}_t) = (s^2_{x,t} + s^2_{x,t+1} - 2r\, s_{x,t}\, s_{x,t+1})/ n'.$$

Muutoksen estimointi on siis tehokkainta käytettäessä täysin samaa näytettä molemmilla jaksoilla.

Näyte tulee helposti harhaiseksi, jos samoja otantayksiköitä rasitetaan tiedusteluilla useita kertoja peräkkäin. Näin saatu näyte on myös täysin joustamaton, ja jos se ei ole alunperin kyllin edustava, ei se täysin samana toistettaessakaan tätä ominaisuutta saa. Tämän vuoksi on monessa tapauksessa päädytty ketjutusratkaisuun, jossa otosyksiköt viipyvät näytteessä vain muutaman tiedustelun verran siten, että määräosa näytteestä uusitaan uutena tiedustelujaksona. Seuraavassa tullaankin tarkemmin käsittelemään tällaisen osittain toistetun otannan ominaisuuksia ja käyttöä hakkuupoistuman määritykseen.

1) Cochran: mts. 336.

5.3. Ketjuotanta

Ketjutusmenetelmää lähdetään seuraavassa yksinkertaisuuden vuoksi tarkastelemaan satunnaisotannan pohjalta.[1] Estimoinnin kohteena ovat mitattavan muuttujan keskiarvot μ_t ja μ_{t+1} ajankohtina t ja t+1 sekä tapahtunutta muutosta kuvaavana näiden keskiarvojen erotus $\Delta = \mu_{t+1} - \mu_t$.

Lähtökohtana on neljä osaotosta:

1) vain ajankohtana t käytetty otos, keskiarvo = $\bar{X}_u$,
2) ajankohtana t käytetty, uudelleen mitattu otos, " = $\bar{X}_m$,
3) ajankohtana t+1 tehty uusi otos, " = $\bar{Y}_n$,
4) kohtaa 2 vastaava otos ajankohtana t+1 " = $\bar{Y}_m$.

Eri ajankohtia vastaavat otosmäärät ovat

$$N_t = m+u \qquad \text{ja} \qquad N_{t+1} = m+n .$$

5.3.1. Tasoestimaattori

Yleisessä muodossa voidaan ajankohdan t+1 tasoestimaatin keskiarvo esittää seuraavasti:

$$\bar{y} = a\,\bar{X}_u + b\,\bar{X}_m + c\,\bar{Y}_m + d\,\bar{Y}_n . \tag{5.5}$$

Koska $E(\bar{X}_u) = E(\bar{X}_m) = \mu_t$ ja $E(\bar{Y}_n) = E(\bar{Y}_m) = \mu_{t+1}$,

täytyy $a + b = 0$ ja $c + d = 1$, joten b voidaan lausua a:n ja d lausua c:n avulla.

$\bar{y}$:n virhevarianssi voidaan muodostaa tavanomaista odotusarvomenettelyä käyttäen, ja tulokseksi saadaan seuraava yhtälö:

---

1) Esitys pohjautuu pääosiltaan artikkeliin Ware and Cunia: Continous Forest Inventory With Partial Replacement of Samples. Forest Science, Monograph 3, Washington D.C. 1962, s. 5-21.

$$(5.6) \qquad V(\bar{y}) = a^2 s_X^2 \left(\frac{1}{u} + \frac{1}{m}\right) + c^2 \frac{s_Y^2}{m} + (1 - c)^2 \frac{s_Y^2}{n}$$

$$- 2\,a\,c\,r \frac{s_X \, s_Y}{m}, \text{ jossa}$$

r on ajanjaksojen t ja t+1 populaation havaintojen välinen korrelaatio, joka voidaan estimoida yhteisestä alanäytteestä m.

Kun a ja c ratkaistaan asettamalla niiden suhteen lasketut osittaisderivaatat nolliksi ja merkitään

$$\bar{Y}_r = \bar{Y}_m + \beta_{XY} (\bar{X} - \bar{X}_m) \text{ , jossa}$$

$$\beta_{XY} = r \frac{s_Y}{s_X} \qquad \text{ja} \qquad \bar{X} = \frac{m}{N_{t+1}}(\bar{X}_m) + \frac{u}{N_{t+1}}(\bar{X}_u) \text{ ,}$$

po. N_t

saadaan $\bar{y}$ lausutuksi muodossa:

$$(5.7) \qquad \bar{y} = c\,(\bar{Y}_r) + (1 - c)\,(\bar{Y}_n) \text{ .}$$

Parametri c määräytyy yhtälöstä

$$c = \frac{\frac{s_Y^2}{n}}{\frac{s_Y^2}{n} + \left[\frac{s_Y^2 (1 - r^2)}{m} + \frac{r^2 s_Y^2}{m+u}\right]}$$

Käyttämällä hyväksi tätä c:n lauseketta voidaan virhevarianssi kirjoittaa lyhyesti

$$(5.8) \qquad V(\bar{y}) = \frac{s_Y^2 \left[1 - (P_u) r^2\right]}{N_{t+1} - (P_u) n r^2} \text{ , jossa} \qquad P_u = u \,/\, N_t \text{ .}$$

5.3.2. Muutosestimaattori

Seuraavaksi tarkastellaan muutoksen $\Delta = \mu_{t+1} - \mu_t$ estimoimista. On mahdollista kehittää useitakin harhattomia estimaattoreita, mutta seuraavassa rajoitutaan vain kaikkein tehokkaimpaan.

Perusolettamukset ovat samat kuin edellä tasoestimaattia käsiteltäessä. Yleisessä muodossa voidaan muutosestimaattori siten lausua analogisesti yhtälön (5.5) kanssa. Samanlaisten käsittelyiden kautta kuin yllä päädytään seuraavaan muutosestimaattoriin ja sen virhevarianssiestimaattoriin:

$$(5.9)\qquad g = \left[\frac{m}{H}(\bar{Y}_r) + \frac{n\left[1 - (P_u)r^2\right]}{H}(\bar{Y}_n)\right]$$

$$- \left[\frac{N_{t+1}(P_m)}{H}(\bar{X}_r) + \frac{(P_u)(N_{t+1} - nr^2)}{H}(\bar{X}_u)\right] \text{, jossa}$$

$$P_m = m / N_t \text{ , } \quad H = N_{t+1} - (P_u)nr^2 \text{ ,}$$

$$\bar{X}_r = \bar{X}_m + \beta_{XY}(\bar{Y} - \bar{Y}_m) \quad \text{ja} \quad \bar{Y} = \frac{m}{N_{t+1}}(\bar{Y}_m) + \frac{n}{N_{t+1}}(\bar{Y}_n) \text{ .}$$

$$(5.10)\qquad V(g) = \left[\frac{1}{N_{t+1} - (P_u)nr^2}\right]$$

$$\cdot\left[\frac{(N_{t+1} - nr^2)}{N_t}(s_X^2) + \left[1 - (P_u)r^2\right](s_Y^2) - 2P_m r s_X s_Y\right].$$

Estimaatti g on yhtälön (5.9) mukaisesti kahden ajanjakson painotettujen keskiarvojen erotus, joten se itse asiassa voidaan lyhyesti kirjoittaa muotoon $g = \bar{y} - \bar{x}$. Tämä ei kuitenkaan tarkoita sitä, että lisäämällä g ajanjakson t tasoestimaattiin $\bar{X}$ saataisiin tulokseksi $\bar{y}$.

5.3.3. Otoksen optimaalinen koko

Alkuperäinen otoskoko sekä toistamiseen mitattujen ja uusien otosyksiköiden määrät riippuvat oleellisesti siitä, onko kysymyksessä taso- vai muutosestimaatti. Tehokkain tasoestimaatti saadaan osittaisella uudelleen mittauksella, kun taas muutosestimaatissa yleensä päästään minimivirheeseen mittaamalla uudestaan kaikki alkuperäiset näyteyksiköt. Käytännössä on usein tyydyttävä kompromissiin, jossa otoskoko määräytyy molempien estimaattien mukaan.

Seuraava ratkaisu perustuu ajankohdan t+1 tehokkaan tasoestimaatin löytämiseen. Ajankohdan t otoskokoa N_t voidaan pitää annettuna. Tarkastelun kohteena on tapaus, jossa keskivirhe on etukäteen määrätty ja kyseessä on siten kustannusten minimointi. Lähtien lineaarisuusoletuksesta on siis tehtävänä minimoida funktio

(5.11) $$C = c_n\, n + c_m\, m + c_f\,, \text{ jossa}$$

c_f edustaa kiinteitä kustannuksia, seuraavan rajoituksen vallitessa

$$V(\bar{y}) \leq V \quad , \quad \sqrt{V} = \text{annettu keskivirhe} .$$

Asettamalla rajoitusyhtälössä yhtäsuuruus voimaan voidaan yhtälöistä (5.11) ja (5.8) muodostaa seuraava Lagrangen funktio:

(5.12) $$\psi = (c_n n + c_m m + c_f) + \lambda\left\{\frac{s_Y^2\left[N_t(1 - r^2) + mr^2\right]}{mN_t + nN_t(1 - r^2) + mnr^2} - V\right\}.$$

Kun tästä yhtälöstä lasketaan osittaisderivaatat n:n, m:n ja kertoimen λ suhteen, asetetaan ne nolliksi ja ratkaistaan näin saadut yhtälöt samanaikaisesti, päädytään seuraaviin lausekkeisiin:

(5.13) $$m = \frac{N_t\sqrt{1 - r^2}}{r^2}\left[\sqrt{\frac{c_n}{c_m}} - \sqrt{1 - r^2}\right],$$

$$(5.14) \qquad n = \frac{s_Y^2}{V} - \frac{N_t}{r^2}\left[1 - \sqrt{\frac{c_m\,(1 - r^2)}{c_n}}\,\right]$$

Uudelleen mitattavien otosyksiköiden määrä ei siis riipu lainkaan asetetusta tarkkuusvaatimuksesta, vaan vain termeistä N_t, r ja c_n/c_m. Lukumäärä m lisääntyy N_t:n tai c_n:n suuretessa, mutta vähenee c_m:n suuretessa.

5.3.4. Ketjuotannan käyttö hakkuupoistuman estimoinnissa

Edellä kuvattua ketjuotantaa on metsäntutkimuksessa käytetty verraten vähän. Ainoa empiirinen tutkimus pohjoismaissa tällä alalla lienee A. Nyyssösen selvitys metsänarvioinnista kahden teollisuusyhtymän metsissä.[1] Mainittu tutkimus on sikäli vajavainen, että uusia otosyksiköitä ei jälkimmäisenä ajankohtana otettu, vaan tyydyttiin pelkkään osaotoksen uudelleen mittaukseen.

Esillä olevassa hakkuupoistuman tapauksessa selostettu otantatapa edellyttää systemaattista otantaa, kun sen sijaan käsitelty ketjuotannan teoria perustui satunnaisotantaan. Tämä heikkous jätetään seuraavassa vaille huomiota, sillä perusjoukko ja otantakehikko ovat sellaiset, että kovin harhaan ei ole mahdollista joutua.

Kaksiulotteiseen otantaan voidaan ketjutusmenettelyä soveltaa siten, että uusia otosyksiköitä varten määrätään aina uutena ajanjaksona oma satunnaislukuparinsa, joka määrää uusien yksiköiden paikan neliöalueessa. Satunnaispisteet on parasta valita palauttamatta valittua lukuparia perusjoukkoon, jolloin eivät samat alueet voi joutua normaaleja toistoja useammin näytteeseen.

Tunnuslukujen ja otoskokojen laskemisessa voidaan käyttää kappaleessa 5.3. esitettyjä estimaattoreita. Koska ne on kehitetty yksinkertaista satunnaisotantaa silmälläpitäen, saattaa olla, että varianssit muodostuvat keskimäärin todellista suuremmiksi. Tämä ei kuitenkaan vaikuta toistettavien rypäiden määrään.

---

1) Nyyssönen: Remeasured sample plots in forest inventory. Meddelelser fra det Norske skogsforsøksvesen, Bind XXII, Oslo 1967, s. 189-220.

Oma mielenkiintonsa on siinä, kuinka usein toisto tulee suorittaa. On todennäköistä, että eri ajankohtien havaintojen välinen korrelaatio pienenee aikavälin kasvaessa. Sen vuoksi on pyrittävä löytämään optimiväli, jolla korrelaatio on vielä tarpeeksi suuri, mutta joka ei ole liian lyhyt uutta mittausta ajatellen. Tavallisimmin kuitenkin käytännön tarpeet sanelevat aikavälin pituuden. Ware ja Cunia ovat saaneet metsien inventoinnin yhteydessä aineistoa, jossa korrelaatiot vaihtelevat eri tutkimuksissa välillä 0.67 - 0.93. Aikaväli on tällöin ollut keskimäärin yhdeksän vuotta. Nyyssönen sai edellä mainitussa tutkimuksessaan arvoja väliltä 0.64 - 0.92 aikavälin ollessa kaikissa tapauksissa viisi vuotta.

Seuraavassa oletetaan, että kustannukset otosyksikköä kohti ovat samat sekä toistetuissa että uusissa otosyksiköissä. Samoin oletetaan, että alkuperäinen otoskoko $N_t = s_Y^2/V$. Tällöin lausekkeet (5.13) ja (5.14) tulevat yhtäsuuriksi eli m=n. Erisuuruisia korrelaatioita saadaan vastaamaan taulukon 4 otantaosuudet. Siinä on $P_m = m/N_t = n/N_t = P_n$ ja $P = m+n/N_t = 2m/N_t$. Viimeisellä rivillä on esitetty tarvittavien näyterypäitten määrät laskettuna N_t:n arvosta 1 317.

r	0.50	0.60	0.70	0.80	0.85	0.90	0.95	0.99
P_m	0.46	0.44	0.42	0.38	0.35	0.30	0.24	0.12
P	0.92	0.88	0.84	0.76	0.70	0.60	0.48	0.24
m+n	1 212	1 159	1 106	1 001	922	790	632	316

Taulukko 4.

Korrelaation täytyy nousta yli 0.5:n ennen kuin sen vaikutus on näkyvä. Vaikka etukäteen ei olekaan voitu laskea korrelaation arvoa hakkuupoistumalle, on silti ilmeistä, että tämä vaatimus on täytetty, koska esimerkiksi kiinteistöjen käyttämän raakapuun korrelaatio peräkkäisinä vuosina on erilaisilla kiinteistotyypeillä välillä 0.~~75~~ 91 - 0.~~91~~ 93.[1]

1) Ervasti, Salo, Seppälä ja Tiililä: Suomen kiinteistöjen raakapuun ja polttoaineiden käytön tutkimus vuosina 1964-66. Metsäntutkimuslaitoksen julkaisuja.

6. YHTEENVETO JA OTANTASUUNNITELMAN TOTEUTUKSEN ETEENPÄIN VIENTI

Kun edellä esitettyä otantasuunnitelmaa tarkastellaan, on sen hyvyyden kriteerinä oltava päätöksentekoon tarvittavan riittävän informaation saaminen käytössä olevien resurssien puitteissa. Päätöksentekijöinä on tällöin nähtävä kokonaisuuden kannalta sen hierarkian yläpäässä olevat elimet. Määrätyn yksityisen metsänomistajan päätöksentekoon tuskin tämänkaltainen suuria kokonaisuuksia käsittävä informaatio vaikuttaisi. Julkisen vallan päätöksentekoa sen sijaan voidaan helpottaa ja nopeuttaa sekä sen toimenpiteiden kautta vaikuttaa puuntuottajiin kokonaisuutena.

Jos tarkastellaan julkisen vallan metsätaloutta koskevaa päätöksentekoa Törnqvistin päätösprosessin valossa[1], voidaan hakkuupoistumasta saatavaa informaatiota yhdessä metsän inventoinnin tulosten kanssa sanoa käytettävän erityisesti tulevaisuuden näkymien muodostamiseen ja niiden arvostamiseen. Myös täytäntöönpanovirheiden ja niiden seurausten havainnoinnissa tarvitaan informaatiota hakkuupoistumasta.

Koska päätöksentekoa varten tarvittavalle tietokokoelmalle on eduksi, että se on mahdollisimman laaja ja moniulotteinen, ei

---

1) Törnqvist esittää päätösprosessin suljettuna systeeminä seuraavasti:

↓
Päätöksentekotilanteen tutkiminen
↓
Mahdollisten vaihtoehtojen löytäminen
↙ ↘
Tulevaisuuden näkymien muodost. | Tulev. näkym. arvostaminen
↓ ↓
Päätöksen valinta eri vaihtoehdoista
↓ PÄÄTÖKSEN TÄYTÄNTÖÖNPANO
Täyt. panovirheiden ja -seurausten havaits.

Törnqvist: Johdatus päätöksentekoteoriaan. Moniste, Helsingin Yliopisto, 1963, s. 6.

tämänkaltaisessa tutkimuksessa kannata tyytyä vain yhden muuttujan mittaamiseen. Hakkuupoistuman yhteydessä tulevat tietysti selvitetyksi sen koostumus monessa eri suunnassa, mutta sen lisäksi kannattaa mitata muitakin tekijöitä, ennenkaikkea sellaisia, joilla voidaan kuvata, selittää tai ennustaa hakkuupoistumaa. Tällöin täytyy otoskoon määräämisen yhteydessä ottaa huomioon näiden muiden tekijöiden estimaateille asetettavat tarkkuusvaatimukset ja niiden mittaamisen aiheuttamat lisäkustannukset.

Tietojen tuoreus ja niiden nopea saanti vaikuttavat olennaisesti päätöksenteon nopeuteen ja hyvyyteen. Suuntaa-antavien päätösten aikaväli määrittää siten myös tutkimusten suorittamisen välisen ajan. Hakkuupoistuman tarkan määrityksen ajankohdan valinnassa tulee ottaa huomioon myös päätöksentekoinformaation toinen metsällinen puoli, volyymin ja kasvun määrittäminen. Tuntuu siten järkevältä keskittää hakkuupoistumatutkimukset metsän inventointien yhteyteen ainakin ajallisesti.

Selostetun otantamenetelmän käytännöllistä toteutusta ajateltaessa on ensimmäisenä edessä esitutkimuksen suorittaminen. Tällöin saadaan selvitys niihin muuttujiin, joille on tähän mennessä täytynyt antaa vain suureksi osaksi harkinnanvaraisia arvoja. Tärkeimpänä tavoitteena on rypäiden optimikoon ja niiden tarvittavan otoslukumäärän selvittäminen. Tähän liittyy olennaisesti vaihtoehtoisten mahdollisuuksien kustannusanalyysi.

Ketjuotantaa ajatellen ei yhtenä ajanjaksona tehty esitutkimus tuo paljon valaistusta, koska kriteerinä on muuttujien välinen korrelaatio ajankohdasta toiseen sekä eri ajanjaksojen kustannusten suhde. Kohdistamalla esitutkimuksen otosyksiköitä osittain samoihin metsälöihin kuin Salon hakkuupoistumatutkimuksessa saadaan näistäkin tekijöistä osviittoja.

LÄHDELUETTELO:

W.G. Cochran	Sampling Techniques, New York 1963.
S. Ervasti, E. Salo, R. Seppälä ja P. Tiililä	Suomen kiinteistöjen raakapuun ja polttoaineiden käytön tutkimus vuosina 1964-66. Metsäntutkimuslaitoksen julkaisuja.
M.H. Hansen, W.N. Hurwitz and W.G. Madow	Sample Survey Methods and Theory, New York 1962.
L. Heikinheimo	Kantorahamalli. Metsätaloudellinen aikakauslehti n:o 12, Helsinki 1966.
V. Holopainen	Suomen metsien luovutusmäärä. Silva Fennica 97, Helsinki 1959.
Y. Ilvessalo	Suomen metsät. Tulokset vuosina 1921-1924 suoritetusta valtakunnan metsien arvioimisesta. Metsätieteellisen koelaitoksen julkaisuja 11, Helsinki 1927.
Y. Ilvessalo	Tutkimuksia yksityismetsien tilasta Hämeen läänin keskiosissa. Acta Forestalia Fennica 26, Helsinki 1923.
H. Järvinen	Tutkimus Suomen postiliikenteestä vuosina 1963-1965. Lisensiaattitutkielma, Helsingin Yliopisto 1966.
M.G. Kendall and W.R.Buckland	A Dictionary of Statistical Terms, London 1966.
A. Langsaeter	Om beregning av middelfeilen ved regelmässige linjetaksering. Meddelelser fra det Norske skogsforsøksvesen, Oslo 1926.
M. Lappi-Seppälä	Linja-arvioimisesta ja sen tarkkuudesta. Metsätieteellisen koelaitoksen julkaisuja 7, Helsinki 1924.
J.W. Lindeberg	Über die Berechnung des Mittelfehlers des Resultates einer Liniertaxierung. Acta Forestalia Fennica 25, Helsinki 1924.
J.W. Lindeberg	Zur Theorie der Liniertaxierung. Acta Forestalia Fennica 31, Helsinki 1927.

B. Matérn	Spatial Variation. Meddelanden från Statens skogsforskningsinstitut 49, Stockholm 1960.
A. Nyyssönen	Remeasured sample plots in forest inventory. Meddelelser fra det Norske skogsforsøksvesen, Bind XXII, Oslo 1967.
E. Salo	Etelä-Pohjanmaan ja Vaasan metsänhoitolautakuntien hakkuupoistuma hakkuuvuonna 1962-63. Julkaisematon tutkimus, Metsäntutkimuslaitos, Metsäekonomian tutkimusosasto.
P. Seeger	Variance Analysis of Complete Designs, Stockholm 1966.
G. Snedecor	Statistical Methods, Iowa 1962.
P.V. Sukhatme	Sampling Theory of Surveys With Applications, Iowa 1954.
L. Törnqvist	Johdatus päätöksentekoteoriaan. Moniste, Helsingin Yliopisto 1963.
L. Törnqvist	The Theory of Replicated Systematic Cluster Sampling With Random Start. Review of ICI, Vol 31:1, 1963.
K.D Ware and T. Cunia	Continous Forest Inventory With Partial Replacement of Samples. Forest Science, Monograph 3, Washington D.C. 1962.

- Proceedings the National Forest Congress (Denver)