

PRO GRADU -TUTKIELMA
Matematiikan, tilastotieteen ja filosofian laitos
Tilastotiede
Marraskuu 2006

Petri Hannuksela

Tukkien tavoite- ja toteutuneiden jakaumien vertailua

Tampereen Yliopisto

Matematiikan, tilastotieteen ja filosofian laitos

HANNUKSELA, PETRI: Tukkien tavoite- ja toteutuneiden jakaumien

vertailua

Pro gradu –tutkielma 41 s., 8 liites.

Tilastotiede

Marraskuu 2006

TIIVISTELMÄ

Suomessa metsät ovat tärkeä luonnonvara. Niistä saadaan esimerkiksi puuta raaka-aineeksi sahateollisuudelle. Sahat tarvitsevat menetelmiä tukkien tavoite- ja toteutuneen halkaisija-pituusjakauman vertailemiseksi, jotta niiden toiminta ja tukkien hyödyntäminen olisi mahdollisimman tehokasta. Tässä tutkielmassa tarkoituksena on esitellä ja vertailla jakaumien yhteensopivuusmittoja sekä soveltaa niitä simuloituun aineistoon.

Tämän tutkielman ensimmäisessä luvussa kerrotaan tukkien katkontamenetelmän taustasta sekä tukkien arvon määrittelystä. Toisessa luvussa esitellään päällekkäisyyskerroimen yhteys jakauma-asteen kanssa. Kolmannessa luvussa käydään läpi kolme yhteensopivuusmittaa, joita voidaan käyttää tukkien tavoitejakauman ja toteutuneen jakauman vertailuun. Nämä yhteensopivuusmitat ovat jakauma-aste, Laspeyresin määräindeksi ja χ^2 -yhteensopivuustesti. Tässä luvussa esitellään myös jakauma-asteen odotusarvo. Neljännessä luvussa esitellään käytetty aineisto, tehdään kolmannessa luvussa esitellyt testit ja tulkitaan saatuja tuloksia. Viidennessä luvussa tehdään simulointi normaalijakauman avulla ja tutkitaan miten yhteensopivuusmittojen tunnusluvut käyttäytyvät varianssin kasvaessa.

Asiasanat: jakauma-aste, päällekkäisyyskerroin, erilaisuusindeksi

Sisältö

1. Johdanto	4
1.1. Tukkien katkontamenetelmän taustaa ja historiaa	4
1.2. Asiakaslähtöisyys sahojen lähtökohdaksi.....	6
1.3. Tukkien arvon määrittely.....	8
2. Päällekkäisyyskerroin ja erilaisuusindeksi.....	10
2.1. Päällekkäisyyskerroimen määritelmä.....	10
2.2. Erilaisuusindeksin esitys ja yhteys päällekkäisyyskerroimen kanssa	11
2.3. Erilaisuusindeksin toinen esitys	12
2.4. Jakauma-asteen ja erilaisuusindeksin yhteys.....	13
3. Yhteensopivuusmitat.....	14
3.1. Matriiseista	14
3.11. Tavoitematriisi	14
3.12. Toteutuneeseen matriisiin vaikuttavat tekijät	15
3.2. Jakauma-aste.....	18
3.21. Jakauma-asteen odotusarvo	19
3.22. Jakauma-asteen varianssi ja luottamusrajat	21
3.3. Hintapainotettu jakauma-aste	22
3.4. Laspeyresin määräindeksi.....	23
3.5. χ^2 -yhteensopivuustesti	24
4. Aineisto ja testit.....	25
4.1. Aineiston kuvaus	25
4.2. Kuusen sahatukkiaineisto	26
4.3. Kuusivaneritukkiaineisto	28
4.4. Kuusen sahatukkiaineiston testitulokset.....	29
4.5. Kuusivaneritukkiaineiston testitulokset.....	30
4.6. Jakauma-asteen odotusarvot	30
4.7. Yhteensopivuusmittojen tulokset leimikoille erikseen.....	31
4.8. Yhteensopivuusmittojen tulosten tulkintaa.....	32
5. Simulointi	35
5.1. Normaalijakauma	35
5.2. Simulointialgoritmi	36
5.3. Simuloinnin tulokset	36
5.4. Simulointituloksien tulkintaa.....	38
Lähteet.....	40
Liite: R-koodit.	42

1. Johdanto

1.1. Tukkien katkontamenetelmän taustaa ja historiaa

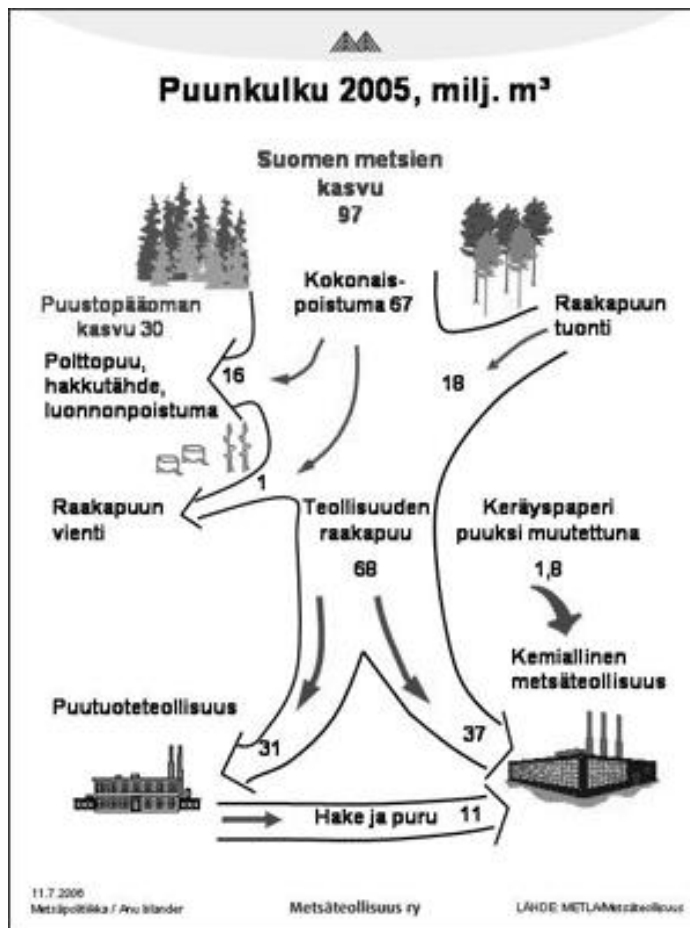
Suomen tärkein luonnonvara ovat metsät, joista saatavaa puuta metsäteollisuus hyödyntää (Kauhanen 1997). Metsät peittävät Suomen pinta-alasta yli kolme neljäsosaa ja metsätalousmaata on kaikkiaan 26 miljoonaa hehtaaria. Suomi kuuluu pääosin pohjoiseen havumetsävyöhykkeeseen. Yleisimmät puulajit täällä ovat mänty, kuusi ja koivu. Puustosta on mäntyä 47 %, kuusta 34 % ja lehtipuuta 19 %. Puuston määrä on lisääntynyt 1970-luvun alusta yli neljänneksen; nykyisin Suomen metsissä on puuta noin 2000 miljoonaa kuutiota. ([http://www.forestindustries.fi/metsa/metsavarat/.](http://www.forestindustries.fi/metsa/metsavarat/)) Myös metsäteollisuuden puunkäyttö on lisääntynyt; se on kaksinkertaistunut viimeisen 40 vuoden aikana (<http://www.forestindustries.fi/metsa/puunkaytto>).

Metsäteollisuuden viennin arvo on suunnilleen kolmasosa Suomen koko viennistä, kun mukaan lasketaan sekä mekaaninen että kemiallinen metsäteollisuus. Metsäteollisuus on siis jakautunut kahteen päähaaraan, mekaaniseen ja kemialliseen. (Kauhanen 1997.)

Kemiallisessa metsäteollisuudessa puuainesta muokataan mekaanisten menetelmien lisäksi erilaisin kemiallisin menetelmin. Kemiallisen metsäteollisuuden piiriin kuuluvat mm. sellu- ja paperiteollisuus. (Kauhanen 1997.)

Mekaanisessa metsäteollisuudessa on kyse puun jalostamisesta mekaanisin keinoin ja sen piiriin kuuluvat esimerkiksi lastulevy- ja sahateollisuus (Kauhanen 1997). Mekaanista metsäteollisuutta kutsutaan myös puutuoteteollisuudeksi (Metsäteollisuus ry 2000). Kuvassa 1.1 kuvataan puunkulku vuodelta 2005 miljoonina kuutiometreinä (<http://www.forestindustries.fi/metsa/puunkaytto/puunkulku2000.html>).

Sahateollisuus on mekaanisen metsäteollisuuden suurin tuotannonala. Suomessa on noin 4 000 erikokoista sahaa, joista suurin osa on pikkusahoja, jotka tuottavat sahatavaraa paikalliseen kulutukseen. Sahateollisuuden tuotteita ovat monet erityyppiset laudat ja lankut sekä eri tavoin käsitellyt jatkojalosteet. Sahateollisuuden tuotteet päätyvät pääasiassa rakennusteollisuuden käyttöön. (Kauhanen 1997.) Sahatavaran osuus on lähes puolet Suomen koko mekaanisen metsäteollisuuden liikevaihdosta. Suomi on Euroopan neljänneksi suurin havusahatavaran tuottaja Venäjän, Ruotsin ja Saksan jälkeen (<http://www.forestindustries.fi/tuotteet/puutuotteet/sahatavara.html>; Metsäteollisuus ry 2000).



Kuva 1.1. Puunkulku Suomessa vuonna 2005.

Apteeraus tarkoittaa rungon katkaisukohtien määrittämistä ottaen huomioon eri puutavaralajien mitta- ja laatuvaatimukset sekä jakaumatavoitteet. Se on puun jalostuksen ensimmäinen ja tärkein vaihe,

koska virheellistä katkontaa on vaikeaa ja kallista korjata tuotantoprosessin myöhemmissä vaiheissa. (Kivinen 2004.)

Puiden rungot voidaan katkoa tavaralajeiksi metsässä ja siitä käytetään nimeä tavaralajimenetelmä (cut-to-length method), tienvarsivarastolla, väliterminaalissa tai käyttöpaikalla (kokopuu-, osapuu-, runko- ja tukkiosamenetelmä). Suomessa runkojen katkaisukohdan merkinnän hoitivat 1960-luvulle asti jakomiehet ja 1960-luvulta alkaen metsurit. 1980-luvulla otettiin käyttöön hakkuukoneita ja 1990-luvulla ns. osarunkoterminaleja. (Kivinen 2004.) Nykyään teollisuuden raakapuu korjataan Suomessa lähes pelkästään tavaralajimenetelmällä (<http://www.forestindustries.fi/metsa/puunkaytto/puunkorjuu.html>).

Hakkuukoneiden apterauksen ohjausjärjestelmä koostuu runkomuodon ennustamisesta, yksittäisen rungon pölkytyksen optimoinnista sekä korjuukohteiden pölkyjen jakauman ohjauksesta. Runkomuodon ennustaminen on apterauksen ohjausjärjestelmän tärkein osa ja sen tavoitteena on ennustaa runkokäyrä koko rungon pituudelle. Ennuste on yleensä 2-5 kertaa rungon tunnetun osan pituus ja sen tarkkuuteen vaikuttavat läpimitan ja pituuden mittaustarkkuus sekä käytetty ennustusmenetelmä (malli). Ennustusmenetelmät ovat joustavia ja oppivia eli ennuste tarkentuu rungon käsittelyn aikana. (Kivinen 2004.)

1.2. Asiakaslähtöisyys sahojen lähtökohdaksi

Sahateollisuudessa, tukkien katkontamenetelmässä, on yleensä ollut tavoitteena tuotantoon sopimattomien tukkien minimointi – tällaisia ovat esimerkiksi mädäntyneet tai epämuodostuneet tukit. Katkontaa on pidetty onnistuneena silloin, kun tällaisten tukkien osuus on ollut mahdollisimman pieni. Tätä menetelmää kutsutaan mitta-analyysiksi. (Kivinen, Uusitalo, Nummi 2005.)

Mitta-analyysi ei ole välttämättä metsänomistajan kannalta paras vaihtoehto. Metsänomistajan kannalta katkenta on tyydyttävä, kun rungot katkotaan mahdollisimman tehokkaasti mahdollisimman arvokkaaksi laaduksi, esimerkiksi sahatukeiksi. (Kivinen et al. 2005.) Tavoitteena on siis maksimoida puunmyyntitulot sekä rungon kauppahinta ja tukkitilavuus (Kivinen 2004). Katkontatuloksen onnistumisen arviointi riippuu siis myös osapuolen näkökulmasta ja tavoitteista. (Kivinen et al. 2005.)

Nykyään monet sahat ovat omaksuneet asiakaslähtöisen strategian tuotannossaan; asiakastilausten vaatimuksia pidetään tärkeämpinä kuin perinteistä kustannusten minimointia. Katkontaproseduurin pitää tuottaa tehtaille riittävä määrä raakamateriaalia sekä tämän lisäksi myös täyttää vaatimukset tukkien pituus-, halkaisija- ja laatujaumien suhteen. Täten katkontatulosten arvioinnin ei pidä perustua ainoastaan mitta-analyysille, vaan myös tavoite- ja toteutuneiden jakaumien vertailulle. (Kivinen et al. 2005.) Myös asiakaslähtöisyyden vuoksi apterauksen (puun runkojen katkontakohtien määrittämisen) ohjaus on entistä tärkeämpää (Kivinen 2004).

Tavoite- ja toteutuneita jakaumia tallennetaan ja analysoidaan kahdesta syystä. Nykyaikaista katkoprosessia metsäkoneella säätelevät hinta- ja tavoitematriisit. Katkontasimulaattoreissa määritellään millä matriisiasetuksilla saadaan suotuisin tulos eri olosuhteisiin metsäkoneen erilaisilla säädöillä. (Kivinen et al. 2005.) Uudet metsäkoneet mittaavat jatkuvasti rungon pituutta sekä läpimittaa ja puun rungot katkotaan pienemmiksi tukeiksi jo hakkuupaikalla (Uusitalo, Kivinen 2004). Ensimmäiseksi siis tarvitaan mittausmenetelmä, jonka avulla voidaan määrittellä mitkä asetukset ovat sopivimpia eri asiakastilauksiin. (Kivinen et al. 2005.)

Toiseksi, katkontatuloksia käytetään kun halutaan vertailla miten eri metsäkoneet ovat onnistuneet tiettyjen tavoitejakaumien saavuttamisessa tai onko puuntoimittajien välillä ollut merkittäviä eroja. Tämän lisäksi nämä

jälkianalyysit tarjoavat tietoa siitä, kuinka säätää katkontaohjeita, jotta saavutettaisiin halutut tukkijakaumat. (Kivinen et al. 2005.)

Tavoite- ja toteutuneen jakauman suhdetta mitataan erilaisilla yhteensopivuusmitoilla. Näitä ovat esimerkiksi jakauma-aste, χ^2 -yhteensopivuustesti ja Laspeyresin määräindeksi. Yhteensopivuusmittoihin palataan tarkemmin myöhemmin ja ne esitellään kolmannessa luvussa.

1.3. Tukkien arvon määrittely

Suomen metsien puulajimäärä ei ole kovin monipuolinen verrattuna useisiin muihin maailman metsäalueisiin, mutta suomalaisen metsän puulajeja on helppo hyödyntää metsäteollisuudessa ja niistä saadaan erinomaisia raaka-aineita moniin eri jalostustarpeisiin. Mekaaninen metsäteollisuus käyttää paljon mäntytukkeja, joita sahataan laudoiksi, lankuiksi, parruiksi ja muuksi sahatavaraksi. Sahateollisuudessa käytetään männyn lisäksi myös kuusitukkeja. (Kauhanen 1997.)

Puiden rungon optimaalisessa katkonnassa täytyy ottaa huomioon, että erityyppisistä tukeista maksetaan erilaisia hintoja. Tukin halkaisijan ja pituuden yhdistelmien haluttavuutta voidaan kuvata esimerkiksi hintamatriisilla tai suhteellisilla arvoilla eri tavaralajien välillä. Eri puutuotteiden välistä suhdetta voidaan kuvailla perushinnoilla. Esimerkiksi männyllä saattaisi olla tietyissä olosuhteissa seuraavanlaiset perushinnat: erikoistukki 400, sahatukki 300, pikkutukki 200 ja massapuu 100. (Uusitalo, Kivinen 2004.)

Massa- eli kuitupuu on massan valmistukseen sopiva puu. Se ei yleensä kelpaa sahattavaksi. (Metsäteollisuus ry 2000.) Pikkutukki on kuitupuun kokoisesta puuaineksesta tehty puutavaralaji (http://www.mhy.fi/keskipohja/puumarkkinat/fi_FI/index/). Sahatukki on nimensä mukaisesti sahattavaksi tarkoitettu tukki (Metsäteollisuus ry

2000). Erikoistukki on arvokkainta ja sitä saatetaan käyttää esimerkiksi huonekaluihin (Uusitalo, Kivinen 2004).

Nykyisin yleisimmissä sahausmenetelmissä tukin pintaosasta tehdään suoraan haketta. Haketettavan osan alta pintapuusta tuotetaan ohuita 19 – 32 mm paksuja sivulautoja. Tukin keskiosasta tuotetaan järeämpiä 32 – 100 mm paksuja sydänpuutavaroita. Valmis sahatavara lajitellaan koon mukaan, tämän jälkeen kuivataan ja lajitellaan laadun mukaan (Metsäteollisuus ry 2000).

Useimmilla puutuotelajeilla on tiettyjä vaatimuksia tukin halkaisijan (*small end diameter, SED*), pituuden sekä laadun suhteen. Jokaisen metsän rakenne näiden suhteen vaikuttaa eniten lopulliseen toteutuneeseen jakaumaan. Siihen vaikuttavat myös mahdolliset kaupalliset rajoitteet puukaupan suhteen sekä perushintataso. (Uusitalo, Kivinen 2004.)

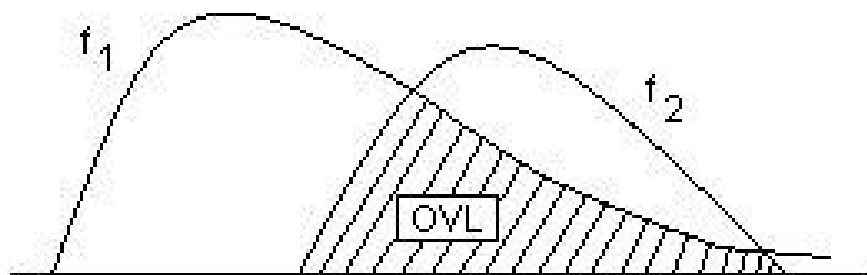
2. Päällekkäisyyskerroin ja erilaisuusindeksi

Käytetyin menetelmä tukkien tavoite- ja toteutuneen jakauman yhteensopivuuden mittaamiseen on jakauma-aste (apportionment degree, apportionment index). Jakauma-aste on tässä luvussa esiteltävän päällekkäisyyskerroimen diskreetti erikoistapaus.

2.1. Päällekkäisyyskerroimen määritelmä

Päällekkäisyyskerroimen yhtäpitävyyden mittana (measure of agreement) esitti alun perin M.S. Weizman vuonna 1970 teoksessaan "Measures of Overlap of Income Distributions of White and Negro Families in the U.S." vertaillen perheiden tulojen (palkkojen) jakaumia. Hän sovelsi päällekkäisyyskerrointa Yhdysvaltojen väestönlaskenta-aineistoon, josta hän vertaili mustien ja valkoisten perheiden tuloja.

Päällekkäisyyskerroin (*overlapping coefficient*, OVL) viittaa kahden todennäköisyys- tai tiheysfunktion yhtäaikaisesti peittämään alueeseen (katso kuva 2.1.). OVL mittaa kahden jakauman yhdenmukaisuutta ja sen



Kuva 2.1. Päällekkäisyyskerroimen graafinen havainnollistaminen.

vaihteluväli on $[0,1]$. Jos $OVL = 0$, jakaumien välillä ei ole päällekkäisyyttä ja jos $OVL = 1$, jakaumat ovat identtiset. (Inman, Bradley 1989.)

Inman ja Bradley (1989) määrittelevät päällekkäisyyskertoimen seuraavalla tavalla. Olkoot $f_1(\underline{X})$ ja $f_2(\underline{X})$ todennäköisyys- tai tiheysfunktioita määriteltynä R_n :ssä. Jatkuvassa tapauksessa OVL määritellään

$$(2.1) \quad \text{OVL} = \int_{R_n} \min [f_1(\underline{X}), f_2(\underline{X})] d\underline{X}.$$

Diskreetissä tapauksessa OVL voidaan määrittellä analogisessa muodossa:

$$\text{OVL} = \sum \min [f_1(\underline{X}), f_2(\underline{X})].$$

Kuvasta 2.1. nähdään, että OVL on se osa todennäköisyysmassasta, joka on yhteinen molemmille jakaumille. Kahden tunnetun jakauman OVL voidaan laskea määrittämällä pisteet, joissa $f_1(\underline{X})$ ja $f_2(\underline{X})$ risteävät ja määrittämällä alueen integraali (yhtälö 2.1).

2.2. Erilaisuusindeksin esitys ja yhteys päällekkäisyyskertoimen kanssa

Inman ja Bradley (1989) esittävät erilaisuusindeksin ja päällekkäisyyskertoimen yhteyden: erilaisuusindeksi on OVL:n komplementti, joka määritellään jatkuvassa tapauksessa

$$D = \frac{1}{2} \int_{R_n} |f_1(\underline{X}) - f_2(\underline{X})| d\underline{X}.$$

Diskreetissä tapauksessa D määritellään

$$D = \frac{1}{2} \sum |f_1(\underline{X}) - f_2(\underline{X})|.$$

Koska $f_1(\underline{X})$ ja $f_2(\underline{X})$ ovat ei-negatiivisia, voidaan kirjoittaa

$$\min[f_1(\underline{X}), f_2(\underline{X})] = \frac{1}{2} [f_1(\underline{X}) + f_2(\underline{X}) - |f_1(\underline{X}) - f_2(\underline{X})|]$$

ja yhtälöstä (2.1) osoittaa, että $OVL = 1 - D$. Täten D vastaa sitä osaa todennäköisyysmassasta, jolle $f_1(\underline{X})$ ja $f_2(\underline{X})$ ei ole yhteistä.

2.3. Erilaisuusindeksin toinen esitys

Sosiologiassa yhteisöllinen erottelu (*social segregation*) on yleinen tutkimuksen aihe. Se tutkii eroja eri ryhmien (esimerkiksi valkoiset ja ei-valkoiset lapset) välillä eri organisaatioyksiköissä (esimerkiksi koulut) eri yhteiskunnissa (esimerkiksi eri maat). Yleisin keino yhteisöllisen erottelun mittaamiseen on erilaisuusindeksi (*dissimilarity index, DI*), jonka tekivät tunnetuksi Duncan & Duncan (1955). (Koskela et al. 2006.)

Erilaisuusindeksin kaava on

$$(2.2) \quad DI = 0.5 * \sum_{i=1}^k |w_i^* - n_i^*|,$$

jossa $w_i^* = \frac{w_i}{\sum_{i=1}^k w_i}$ ja $n_i^* = \frac{n_i}{\sum_{i=1}^k n_i}$ ja w_i^* ja n_i^* ovat esimerkiksi valkoisten

ja ei-valkoisten lapsien lukumäärät koulussa i , $i = 1, \dots, k$ ja $\sum_{i=1}^k w_i$ ja

$\sum_{i=1}^k n_i$ merkitsevät esimerkiksi valkoisten ja ei-valkoisten lapsien

kokonaislukumäärää maassa. Erilaisuusindeksin vaihteluväli on $[0,1]$, jossa arvo nolla tarkoittaa, ettei erottelua ole ja arvo yksi merkitsee täydellistä erottelua (Koskela et al. 2006).

2.4. Jakauma-asteen ja erilaisuusindeksin yhteys

Jakauma-astetta käsitellään luvussa 3.2 ja se määritellään yhtälössä (3.1). Tukkien katkontatulosten vertailu tavoite- ja toteutuneen matriisin avulla on hyvin samankaltaista kuin yhteisöllisen erottelun mittaaminen. Yhteisöllisen erottelun mittaamisessa tosin molempia jakaumia (esimerkiksi valkoiset ja ei-valkoiset lapset) pidetään satunnaisina, kun taas katkontatulosten analysoinnissa vain tukkien toteutunut jakauma on satunnainen. Tavoitematriisi on sahanomistajien tai asiakkaiden antama. (Koskela et al. 2006.)

Jakauma-asteen ja erilaisuusindeksin kaavojen yhteys on selvä,

$$AI = 1 - DI.$$

Kuitenkin pitää huomata, että koska tavoitejakauma on annettu jakauma-asteelle, sen vaihteluväli ei ole sama kuin erilaisuusindeksillä $([0,1])$, vaan $[\min(t_1^*, t_2^*, \dots, t_k^*), 1]$ (luku 3.2). Huolimatta tästä erosta, on näiden kahden käsitteen yhteys huomattava. (Koskela et al. 2006.)

3. Yhteensopivuusmitat

Tavoite- ja toteutuneen jakauman suhdetta voidaan mitata erilaisilla yhteensopivuusmitoilla. Tässä luvussa esitellään jakauma-aste, hintapainotettu jakauma-aste, χ^2 -yhteensopivuustesti sekä Laspeyresin määräindeksi. Ensiksi tutustutaan kuitenkin matriiseihin, joita tarvitaan yhteensopivuusmittojen soveltamisessa. Nämä matriisit ovat tavoite-, hinta- sekä toteutunut matriisi.

3.1. Matriiseista

3.11. Tavoitematriisi

Tavoitejakauma (*demand matrix, target matrix, T*) tukeille ilmaistaan yleensä matriisimuodossa. Tavoitematriisilla $T = (t_{ij})$ tietylle tukkityypille on m riviä (rivi-indeksi $i = 1, \dots, m$) ja n saraketta (sarakeindeksi $j = 1, \dots, n$). Jokainen rivi esittää tiettyä halkaisija (*small end diameter, SED*) – luokkaa ja jokainen sarake viittaa tiettyyn pituusluokkaan. Kivinen, Uusitalo ja Nummi (2005) esittävät tavoitematriisin muodossa:

$$T = \begin{pmatrix} t_{11} & t_{12} & \dots & t_{1n} \\ t_{21} & t_{22} & \dots & t_{2n} \\ \vdots & \vdots & & \vdots \\ t_{m1} & t_{m2} & & t_{mn} \end{pmatrix} = (t_{ij})$$

jossa t_{ij} on tukkien määrä i :nnessä halkaisijaluokassa ja j :nnessä pituusluokassa tukkien tavoitejakaumassa, $i = 1, \dots, m$ ja $j = 1, \dots, n$. Nyt tukki, jonka halkaisija on d ja pituus l kuuluu tukkiluokkaan (i, j) , kun $d_i \leq d < d_{i+1}$ ja $l_j \leq l < l_{j+1}$.

Tavoitematriisiin määrittelyyn on kaksi päävaihtoehtoa. Ensimmäinen tapa on määrittää jokaisen tukkiluokan halutun tukkimäärän suhde

tehtaalle tulevaan tietyn ajan kuluessa tulevaan kokonaistukkimäärään. Jos esimerkiksi tukkien osuus toisessa halkaisija- ($i = 2$) ja kolmannessa pituusluokassa ($j = 3$) on 5%, luku t_{23} valitaan niin, että t_{23} :n suhde kokonaistukkimäärään ($\sum t_{ij}$) on 0.05. Toinen tapa määrittää tavoitematriisi, on esittää se yksinkertaisesti absoluuttisina tukkien määrinä. Tavoitematriisi voidaan aina helposti muuttaa suhteellisesta absoluuttiseksi ja toisin päin. (Kivinen et al. 2005.)

Tukkeja, joilla on suuri halkaisija, ei voida tuottaa pienikokoisista rungoista. Tästä syystä hakkuupaikalla on vain pieni mahdollisuus vaikuttaa tukkien halkaisijajakaumaan. Monet sahat Skandinaviassa ovatkin omaksuneet tavoitematriisityypin, joka määrittelee tavoitteet ainoastaan tukkien pituusjakaumalle. Jos esimerkiksi tavoitematriisin lukujen t_{2j} toisen rivin kumulatiivinen summa on 200 ja toisen pituusluokan tavoiteosuus toisessa halkaisijaluokassa on 5%, niin solussa (2,2) tavoitematriisissa pitäisi olla arvo 10 ($t_{22} = 10$). (Kivinen et al. 2005.)

Vastaavalla tavalla tukeille määritellään hintamatriisi (*price matrix*), merkitsemällä jokaiseen halkaisija- ja pituusluokkaan se hinta, joka ko. luokasta maksetaan. Toteutuneeseen matriisiin (*output matrix*) merkitään tukkien toteutunut pituus- ja halkaisijajakauma puiden korjuun jälkeen.

3.12. Toteutuneeseen matriisiin vaikuttavat tekijät

Uusitalon ja Kivisen (2004) mukaan toteutuneeseen matriisiin vaikuttavat tekijät voidaan jakaa kahteen ryhmään: tekijöihin, joita voidaan simuloida sekä tekijöihin, joita ei voida simuloida. Nämä taulukossa 3.1. näkyvät tekijät esitellään tässä luvussa.

Hintamatriisi muodostaa katkontaoptimoinnin perustan. Hintamatriisin tehtävä on kertoa tietokoneelle kuinka priorisoida erilaisten tukkikokojen katkonta ja näin maksimoida vaaditun tuotteen tuotto. Hintamatriisin yksittäisten solujen arvoilla on suurin vaikutus katkontatulokseen. Tämän

lisäksi hintamatriisin koolla eli solujen lukumäärällä on tärkeä merkitys. Mitä yksinkertaisempi matriisi, sitä helpommin saadaan yksinkertainen matemaattinen ratkaisu ja suurimmassa osassa tapauksia hyvät yhteensopivuusarvot. Myös hintamatriisin arvojen vaihteluvälillä on vaikutusta lopputulokseen; isommalla vaihteluvälillä löydetään todennäköisemmin enemmän toteutettavissa olevia katkontamalleja. (Uusitalo, Kivinen 2004.)

Taulukko 3.1. Toteutuneeseen matriisiin vaikuttavat tekijät.

Voidaan simuloida	Ei voida simuloida
tavoitematriisi	metsäkoneen ajajan vaikutus
hintamatriisi	mittausvirhe
puuvalikoimien suhde	runkokäyrien ennustaminen
aseman muodostaminen	
katkonta-algoritmi	

Kuten hintamatriisin tapauksessa, tavoitematriisin solujen lukumäärä ja solujen arvojen vaihteluväli sekä yksittäisten soluarvojen jakauma vaikuttavat toteutuneeseen jakaumaan. Monilla sahoilla on erilainen lähestymistapa tavoitematriisin luomiseen. Esimerkiksi joillakin sahoilla tavoitematriisi on sovitettu sen tarjonta-alueen metsien ”luonnolliseen puuvirtaan”. Koska toteutunut matriisi harvoin vastaa tavoitejakaumaa täydellisesti, on päätelty että on parempi ohjata jakaumaa siihen suuntaan, joka saavutetaan luonnollisesti. (Uusitalo, Kivinen 2004.)

Suomessa tukkien arvot hintamatriiseissa ovat yleensä suhteellisia arvoja sattumanvaraisesti valitun pisteen ympärillä. Jotta saataisiin paras mahdollinen tulos, puulajitelmien keskimääräisten hintojen suhteiden pitää olla mahdollisimman todenmukaisia. Monesti erot lajitelmien perushintojen välillä on asetettu niin, ettei päällekkäisyyttä eri lajitelmien välillä esiinny. Jos tietystä lajitelmasta on puutetta, pitää sen (suhteellista) hintaa nostaa. Jos tiettyjen lajitelmien arvot ovat lähellä toisiaan tai jopa päällekkäisiä, sillä voi olla ratkaiseva vaikutus toteutuneeseen matriisiin. (Uusitalo, Kivinen 2004.)

Leimikon rakenteella (stand composition) on merkittävä vaikutus katkontatulokseen. Oma vaikutuksensa on puiden lukumäärällä, niiden koolla sekä tukkijakauman muodolla. Jakauma-apteerauksessa (bucking-to demand optimization), hintamatriisia sovitetaan koko ajan ottaen huomioon jo saavutettu kysyntä eri kokojen osalta. Eli mitä enemmän puita leimikossa on, sitä paremmat yhteensopivuusluvut saavutetaan. Mitä isompia puut taas ovat, sitä enemmän on katkontavaihtoehtoja optimoinnissa. Esimerkiksi leimikko, jossa on vain pienikokoisia puita, johtaa huonoihin yhteensopivuuslukuihin, koska suurikokoisten tukkien kysyntää ei saada täytettyä. (Uusitalo, Kivinen 2004.)

Puun rungon katkontaan tukeiksi mahdollisimman optimaalisesti on käytetty dynaamista ohjelmointia (dynamic programming, DP). Sitä sovelsivat tähän ensimmäistä kertaa Pnevmticos ja Mann (1972). Katkonta-algoritmien (bucking algorithm) vaikutusta on vaikea arvioida, koska metsäkoneiden valmistajat eivät yleensä paljasta omia algoritmejaan. (Uusitalo, Kivinen 2004.)

Puun tuntemattoman osan runkokäyrien ennustaminen voidaan toteuttaa useilla eri metodeilla kuten esimerkiksi lineaarisilla sekamalleilla. Ennustusmallilla on merkittävä vaikutus katkontan lopputulokseen. (Uusitalo, Kivinen 2004.)

Pituus- ja halkaisijamittaukset perustuvat mekaaniseen kontaktiin rungon ja metsäkoneen päähän asennettujen sensoreiden välillä. Mittaus tietystä kohdasta vaikuttaa sekä rungon profiiliin että seuraaviin katkontapäätöksiin, johon rungosta kerätty data vaikuttaa. Pituus- ja halkaisijamittauksilla on suuri vaikutus rungon profiilin ennustuksen tarkkuuteen. Täten mahdollisilla mittausvirheillä (measurement error) on vaikutusta lopputulokseen. (Uusitalo, Kivinen 2004.)

Kun metsäkone saavuttaa ennalta valitun pituuden rungossa, voi sen kuljettaja (metsäkoneen ohjaaja, harvester operator) joko antaa katkaisukäskyn tai valita toiseksi parhaan ratkaisun. Käytännössä

metsäkoneen kuljettajat joutuvat joskus hylkäämään metsäkoneen valitseman optimaalisen katkaisukohdan erilaisten tukin laatuluokkaan vaikuttavien rungon vikojen takia. Jokainen kuljettaja tarkkailee, analysoi ja ratkaisee nämä tapaukset omalla tavallaan. Ero toteutuneessa jakaumassa voi olla merkittävä, vaikka kaksi kuljettajaa työskentelisi samanlaisissa olosuhteissa samanlaisilla hinta- ja tavoitematriiseilla. (Uusitalo, Kivinen 2004.)

3.2. Jakauma-aste

Jakauma-aste (*apportionment index (AI)*, *apportionment degree*) on tukkien tavoite- ja toteutuneen jakauman suhteen mittaamiseen käytetyin mitta Skandinaviassa. Sen kehitti ruotsalainen matemaatikko Bergstrand 1980-luvun puolivälissä. Jakauma-aste on saavuttanut suosiota, koska se on yksinkertainen, sitä on helppo käyttää ja tulkita. (Koskela, Sinha, Nummi 2006.) Kivinen, Uusitalo ja Nummi (2005) määrittelevät jakauma-asteen seuraavalla tavalla:

$$(3.1) \quad AI = (1 - 0.5 * \sum_{i=1}^m \sum_{j=1}^n |f_{ij}^* - t_{ij}^*|),$$

jossa m on halkaisijaluokkien lukumäärä, n on pituusluokkien määrä,

$$f_{ij}^* = \frac{f_{ij}}{\sum_{i=1}^m \sum_{j=1}^n f_{ij}}, \text{ jossa } f_{ij} \text{ on tukkien lukumäärä } i\text{:nessä halkaisija- ja}$$

$$j\text{:nessä pituusluokassa toteutuneessa matriisissa ja } t_{ij}^* = \frac{t_{ij}}{\sum_{i=1}^m \sum_{j=1}^n t_{ij}}, \text{ jossa}$$

t_{ij} on tukkien lukumäärä i :nessä halkaisija- ja j :nessä pituusluokassa tavoitematriisissa, $i = 1, \dots, m$ ja $j = 1, \dots, n$. Yhtälö (3.1) voidaan kirjoittaa muodossa

$$(3.2) \quad AI = \sum_{i=1}^m \sum_{j=1}^n \min(f_{ij}^*, t_{ij}^*).$$

Yhtälö (3.1) osoittaa, kuinka jakauma-aste lasketaan, kun tukkien tavoiteosuudet on annettu erikseen jokaiselle halkaisijapituusyhdistelmälle (Kivinen et al. 2005). Jakauma-asteen maksimiarvo on 1 (100%), joka saavutetaan, kun tavoite- ja toteutunut jakauma ovat identtisiä. Minimiarvo on $[\min(t_1^*, t_2^*, \dots, t_k^*), 1]$, joka selvästi riippuu pienimmästä kohdematriisin solusta. Täten jakauma-aste saavuttaa miniminsä, kun kaikki tukit ovat sattuneet pituus-läpimittaluokkaan, jolla on pienin tavoiteosuus (katso yhtälö (3.2)). (Koskela et al. 2006.)

3.21. Jakauma-asteen odotusarvo

Nummi, Koskela ja Sinha käsittelevät artikkelissaan "Some aspects of the sampling distribution of the apportionment index and related interference" (2006) mm. jakauma-asteen odotusarvoa ja varianssia. Tässä alaluvussa esittelen jakauma-asteen odotusarvon tuon esityksen pohjalta.

Oletetaan aluksi, että tukkien toteutunut jakauma noudattaa multinomijakaumaa

$$(f_1, f_2, \dots, f_k) \sim MN(N; t_1^*, t_2^*, \dots, t_k^*),$$

jossa (f_1, f_2, \dots, f_k) on satunnainen toteutunut jakauma, $N = \sum_{i=1}^k f_i$ on

kaadettujen puiden kokonaismäärä ja t_i^* :t ovat suhteelliset tavoitematriisin

arvot, $t_i^* \in (0,1)$ ja $\sum_{i=1}^k t_i^* = 1$.

Suhteellisesta satunnaisesta toteutuneesta jakaumasta käytetään merkintää $f^* = (f_1^*, f_2^*, \dots, f_k^*)$, jossa $f_i^* = \frac{f_i}{N} \in (0,1)$ ja $\sum_{i=1}^k f_i^* = 1$. Tästä seuraa, että

$$f_i \sim \text{Bin}(N, t_i^*)$$

ja edelleen

$$E(f_i^*) = E\left(\frac{f_i}{N}\right) = t_i^*, \quad V(f_i^*) = V\left(\frac{f_i}{N}\right) = \frac{t_i^*(1-t_i^*)}{N} = \sigma_i^2$$

ja

$$\text{Corr}(f_i^*, f_j^*) = -\sqrt{\frac{t_i^* t_j^*}{(1-t_i^*)(1-t_j^*)}}.$$

Binomijakaumaa voidaan tietyissä olosuhteissa approksimoida normaalijakauman avulla. Erään säännön mukaan, kun $X \sim \text{Bin}(n,p)$, pitää $\min(np, (n(1-p))) \geq 5$. Tämä ehto toteutuu käytännössä hyvin, koska tukkien kokonaismäärä (N) on yleensä hyvin suuri. Normaaliapproksimaatio antaa tuloksen

$$\frac{\sqrt{N}(f_i^* - t_i^*)}{\sqrt{t_i^*(1-t_i^*)}} \sim N(0,1).$$

Seuraavaksi oletetaan, että $X \sim N(\theta, \sigma^2)$. Tällöin

$$E(|X - \theta|) = \sigma \sqrt{\frac{2}{\pi}},$$

$$V(|X - \theta|) = \sigma^2 \left(1 - \frac{2}{\pi}\right).$$

Jakauma-asteen odotusarvo voidaan kirjoittaa muodossa (Nummi, Koskela, Sinha (2006))

$$(3.22) \quad E(AI) = 1 - 0.5 * \sum_{i=1}^k E |o_i^* - t_i^*| \\ = 1 - \frac{1}{2} \sqrt{\frac{2}{N\pi}} \sum_{i=1}^k \sqrt{t_i^*(1-t_i^*)}.$$

Erikoistapauksessa, kun $t^* = (\frac{1}{k}, \dots, \frac{1}{k})$, $E(AI)$ sievenee muotoon

$$E(AI) = 1 - \sqrt{\frac{1}{2\pi N}} (k-1).$$

Kun tukkien määrä N kasvaa, jakauma-asteen odotusarvo kasvaa ja $E(AI) \rightarrow 1$ kun $N \rightarrow \infty$. Täten multinomijakaumaoletuksella voidaan odottaa ykköstä lähellä olevia odotusarvoja, silloin kun tukkien määrä N on hyvin suuri. (Koskela et al. 2006.)

3.22. Jakauma-asteen varianssi ja luottamusrajat

Tässä alaluvussa esitellään jakauma-asteen varianssin ja luottamusrajan kaavat Koskelan, Sinhan ja Nummen (2006) esityksen pohjalta. Jakauma-asteen varianssin kaava on

$$V(AI) = \sum_{i=1}^k \frac{t_i^*(1-t_i^*)}{4N} - \frac{1}{2\pi} \left(\sum_{i=1}^k \sqrt{\frac{t_i^*(1-t_i^*)}{4N}} \right)^2 \\ + \sum \sum_{i \neq j} \frac{\sqrt{t_i^* t_j^* (1-t_i^*) (1-t_j^*)}}{4N} \phi \left(\sqrt{\frac{t_i^* t_j^*}{(1-t_i^*) (1-t_j^*)}} \right),$$

jossa $\varphi(\rho) = \frac{2}{\pi} |\rho| \arcsin |\rho| + \frac{2}{\pi} \sqrt{1-\rho^2}$ ja muut merkinnät samat kuin edellisessä alaluvussa. Jakauma-asteen $100 * (1-\alpha) \%$ alempi luottamusraja lasketaan kaavalla

$$E(AI) - z_{\alpha} \sqrt{V(AI)}.$$

3.3. Hintapainotettu jakauma-aste

Jakauma-asteen on kritisoitu olevan liian karkea mitta, koska se antaa saman painoarvon kaikille tukkiluokille (Koskela et al. 2006). Tästä syystä siitä on kehitetty hintapainotettu versio (*price-weighted apportionment degree, AP*). Kivinen, Uusitalo ja Nummi (2005) esittävät sen muodossa

$$(3.3) AP = \sum_{i=1}^m \sum_{j=1}^n \omega_{ij} \min(f_{ij}^*, t_{ij}^*) = \left(\frac{\bar{X}(\omega) + \bar{\theta}(\omega)}{2} - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n \omega_{ij} |f_{ij}^* - t_{ij}^*| \right),$$

jossa $\bar{X}(\omega) = \sum_{i=1}^m \sum_{j=1}^n f_{ij}^* \omega_{ij}$, $\bar{\theta}(\omega) = \sum_{i=1}^m \sum_{j=1}^n t_{ij}^* \omega_{ij}$ ja $\omega_{ij} = \frac{p_{ij}}{\sum_{i=1}^m \sum_{j=1}^n p_{ij}}$.

Ongelmana on, että yhtälö (3.3) ei ole välttämättä mielekkäästi tulkittavissa, koska se on yksilöllisten jakauma-asteiden painotettu keskiarvo yli koko matriisin. Jotta yhtälöstä (3.3) saataisiin järkevä yhteensopivuusmitta, Kivinen, Uusitalo ja Nummi (2005) suhteuttavat AP:n sen maksimiarvoon, joka on

$$(3.4) AP_{MAX} = \sum_{i=1}^m \sum_{j=1}^n t_{ij}^* \omega_{ij}.$$

3.4. Laspeyresin määräindeksi

Taloustieteessä hinnan tai määrän muutosta kuvataan indekseillä. Esimerkiksi kuluttajahintaindeksi (*consumer price index, CPI*) kuvaa keskimääräisen kuluttajan ostamien tavaroiden ja palveluiden hinnan kokonaismuutosta. Normaalisti muutokset hinnassa ja määrässä ovat korreloituneet eli hyödykkeen hinnan muutos vaikuttaa sen kulutukseen. Jotta voitaisiin tunnistaa se hinnan osan muutos, joka johtuu todellisesta hinnan muutoksesta ja se osa, joka johtuu rakenteellisesta muutoksesta kuluttajien käyttäytymisessä, on kehitetty erilaisia painotettuja hinta- ja määräindeksejä. Laspeyresin määräindeksi (*Laspeyres' quantity index, Q_L*) on esimerkki tällaisesta indeksistä, jossa hyödykkeiden hintoja käytetään painotekijöinä. (Kivinen et al. 2005.)

Kun verrataan tukkien toteutunutta jakaumaa tavoitejakaumaan, Laspeyresin määräindeksi voidaan esittää Kivisen, Uusitalon ja Nummen (2005) mukaan seuraavalla tavalla:

$$(3.5) \quad Q_L = \frac{\sum_{i=1}^m \sum_{j=1}^n p_{ij} o_{ij}}{\sum_{i=1}^m \sum_{j=1}^n p_{ij} e_{ij}},$$

jossa p_{ij} kuvaa tukin arvoa tai hintaa i :nessä halkaisijaluokassa ja j :nessä pituusluokassa, o_{ij} on havaittu frekvenssi ja e_{ij} tavoitefrekvenssi tukeille luokassa i, j .

Laspeyresin määräindeksi kuvaa arvojen suhdetta ennen puiden korjuuta ja puiden korjuun jälkeen. Jakauma-asteesta poiketen, Laspeyresin määräindeksi saattaa ylittää arvon 1 (100%). Toisaalta, jos todellinen toteutunut jakauma vastaa tukkien määrässä tavoitejakaumaa, Laspeyresin määräindeksi antaa arvon 1. (Kivinen et al. 2005.)

3.5. χ^2 -yhteensopivuustesti

χ^2 -yhteensopivuustestiä (Pearsonin χ^2 -testi) käytetään yleisesti tilastotieteessä mittaamaan kuinka hyvin havaittu jakauma sopii teoreettiseen jakaumaan. Kun sitä sovelletaan tukkien tavoite- ja toteutuneen jakauman yhteensopivuuden testaamiseen, se voidaan määritellä

$$(3.6) \quad \chi^2 = \sum_{i=1}^m \sum_{j=1}^n \left(\frac{(o_{ij} - e_{ij})^2}{e_{ij}} \right),$$

jossa o_{ij} on havaittu frekvenssi ja e_{ij} tavoitefrekvenssi tukeille luokassa i, j ja joka tietyissä olosuhteissa noudattaa χ^2 -jakaumaa $(nm - 1)$:llä vapausasteella. (Kivinen et al. 2005; Nummi, Sinha, Koskela 2006.)

χ^2 -tunnusluku antaa arvon nolla, jos tavoitematriisi on identtinen toteutuneen matriisin kanssa. Testi kuitenkin tuottaa suuria positiivisia arvoja, jos tavoite- ja toteutuneen jakauman välillä on suuria eroja. Käyttämällä kontingenssikerrointa (*contingency coefficient*, C), voidaan testin arvot skaalata sopivalle vaihteluvälille:

$$(3.7) \quad C = \sqrt{\chi^2 / (\chi^2 + N)},$$

jossa χ^2 on χ^2 -testisuureen arvo ja N on tukkien kokonaislukumäärä. Vähentämällä kontingenssikerroin luvusta yksi $(1-C)$, saadaan käytännöllinen mitta tukkien tavoite- ja toteutuneen jakauman välisen suhteen tutkimiseen. Jos matriisit vastaavat toisiaan täydellisesti, $1 - C$ antaa arvon yksi, muutoin $1 - C$ vähenee kohti nollaa. (Kivinen et al. 2005.)

4. Aineisto ja testit

4.1. Aineiston kuvaus

Tässä tutkielmassa käyttämäni aineisto on saatu Veli-Pekka Kiviseltä. Aineisto sisältää tavoitejakaumat kahdelle kuusipuutavaralajille, normaalille kuusen sahatukille sekä kuusivaneritukille, sekä vastaavien tavaralajien toteutuneet jakaumat 15 eteläsuomalaisessa kuusileimikossa. Seuraavissa alaluvuissa käytetyissä jakaumissa on ensin esitetty tulokset, joissa kaikkien 15 leimikoiden jakaumat on laskettu yhteen. Lisäksi jokaisen 15 leimikon tulokset esitetään myös erikseen.

Toteutuneet jakaumat ovat tuloksia Kivisen simulointiajoista hänen lähioptimaalimenetelmää soveltavalla katkonta-algoritmillaan. Simuloinneissa on käytetty tasahintamatriiseja (50 €/ m³ sahatukki, 65 €/ m³ kuusivaneritukki) sekä hakkuukoneiden keräämää runkodataa. Kullakin rungolla toteutettavaksi on valittu se niistä 10 tutkitusta pölkytysvaihtoehdosta, jonka kokonaisprioriteetti on suurin eli runko on katkottu niihin pituus-latvaläpimitta-yhdistelmiin, joilla on ollut suurin kysyntä.

Jakaumissa läpimitta- ja pituusluokkien arvot ovat luokkien alarajoja: esimerkiksi läpimitta-luokkaan 280 kuuluvat ne tukkikappaleet, joiden latvaläpimitta on välillä 280-299, jos seuraavan luokan alaraja on esimerkiksi 300.

4.2. Kuusen sahatukkiaineisto

Taulukko 4.1. Sahatukit, tavoitematriisi.

SAHATUKKI	370	430	460	490	520	550	580	610	Yhteensä
160	2	8	8	19	8	15	8	8	76
170	2	8	8	19	8	15	8	8	76
180	3	8	8	19	12	15	4	8	77
200	3	15	8	15	8	15	4	8	76
220	0	10	10	19	10	19	4	5	77
240	0	8	9	14	15	19	4	8	77
260	0	8	8	12	19	19	4	8	78
280	0	8	8	12	19	19	4	8	78
300	0	8	8	12	19	19	4	8	78
320	0	9	9	11	15	15	9	8	76
340	0	8	15	19	19	8	4	4	77
360	0	8	15	19	19	8	4	4	77
380	0	8	15	19	19	8	4	4	77
400	0	0	0	0	0	0	0	0	0
Yhteensä	10	114	129	209	190	194	65	89	1000

Taulukko 4.2. Sahatukit, toteutunut matriisi.

SAHATUKKI	370	430	460	490	520	550	580	610	Yhteensä
160	1127	663	730	739	756	693	823	2469	8000
170	86	123	108	286	73	181	95	117	1069
180	1051	102	41	148	46	82	16	67	1553
200	1881	525	199	233	116	210	48	85	3297
220	0	2215	396	352	188	293	105	82	3631
240	0	1133	235	147	117	165	20	41	1858
260	0	360	34	35	90	64	1	1	585
280	0	1	0	0	3	0	0	0	4
300	0	0	0	0	0	0	0	0	0
320	0	0	0	0	0	0	0	0	0
340	0	0	0	0	0	0	0	0	0
360	0	0	0	0	0	0	0	0	0
380	0	0	0	0	0	0	0	0	0
400	0	0	0	0	0	0	0	0	0
Yhteensä	4145	5122	1743	1940	1389	1688	1108	2862	19997

Yllä olevissa taulukoissa on esitetty sahatukkien tavoitematriisi (taulukko 4.1) ja toteutunut matriisi (taulukko 4.2). Sahatukin läpimitta on

esitetty taulukon vasemmalla pystyriivillä millimetreinä ja sen vaihteluväli on 160 – 400 mm. Ylimmällä vaakariivillä taulukossa on sahatukin pituus senttimetreinä, siinä vaihteluvälinä on 370 – 610 cm. Taulukon sisäsolut näyttävät läpimitta-pituusyhdistelmien määrän.

Taulukko 4.3. Sahatukit, hintamatriisi.

SAHATUKKI	370	430	460	490	520	550	580	610
160	3,72	4,32	4,62	4,93	5,23	5,53	5,83	6,13
170	4,20	4,88	5,22	5,56	5,90	6,24	6,58	6,92
180	4,71	5,47	5,85	6,23	6,62	7,00	7,38	7,76
200	5,81	6,75	7,23	7,70	8,17	8,64	9,11	9,58
220	7,03	8,17	8,74	9,31	9,88	10,45	11,02	11,59
240	8,37	9,73	10,40	11,08	11,76	12,44	13,12	13,80
260	9,82	11,41	12,21	13,01	13,80	14,60	15,40	16,19
280	11,39	13,24	14,16	15,09	16,01	16,93	17,86	18,78
300	13,08	15,20	16,26	17,32	18,38	19,44	20,50	21,56
320	14,88	17,29	18,50	19,70	20,91	22,12	23,32	24,53
340	16,80	19,52	20,88	22,24	23,61	24,97	26,33	27,69
360	18,83	21,88	23,41	24,94	26,46	27,99	29,52	31,05
380	20,98	24,38	26,08	27,79	29,49	31,19	32,89	34,59
400	23,25	27,02	28,90	30,79	32,67	34,56	36,44	38,33

Sahatukkien hintamatriisi (taulukko 4.3) on yksinkertaisesti muodostettu laskemalla jokaisen tukin tilavuus neliömetreinä ja kertomalla tulos sahatukin hinnalla, joksi on siis oletettu 50 €/ m³. Taulukossa 4.3 on läpimitta- ja pituusmittayksiköt muutettu metreiksi laskemisen helpottamiseksi. Tukin tilavuus on laskettu kaavalla

$$(4.1.) \quad V = (d / 2)^2 * \pi * h * p,$$

jossa d = tukin läpimitta, h = tukin pituus ja p = tukin hinta (MAOL-tilaukset 1995).

4.3. Kuusivaneritukkiaineisto

Taulukko 4.4. Kuusivaneritukit, tavoitematriisi.

kuusivaneri	330	410	490	530	610	Yhteensä
280	10	10	40	10	10	80
300	10	10	40	10	10	80
320	10	10	40	10	10	80
340	10	10	40	10	10	80
360	10	10	40	10	10	80
380	10	10	60	10	10	100
400	10	10	60	10	10	100
450	10	10	60	10	10	100
500	10	10	60	10	10	100
550	10	10	60	10	10	100
600	10	10	60	10	10	100
650	0	0	0	0	0	0
Yhteensä	110	110	560	110	110	1000

Taulukko 4.5. Kuusivaneritukit, toteutunut matriisi.

Kuusivaneri	330	410	490	530	610	Yhteensä
280	568	458	518	461	642	2647
300	354	52	40	9	7	462
320	436	132	93	39	9	709
340	265	74	78	26	19	462
360	167	31	35	18	9	260
380	142	36	28	11	4	221
400	102	15	26	7	4	154
450	16	2	9	0	2	29
500	2	1	3	0	0	6
550	0	0	0	0	0	0
600	0	0	0	0	0	0
650	0	0	0	0	0	0
Yhteensä	2052	801	830	571	696	4950

Yllä olevissa taulukoissa on kuusivaneritukin tavoitematriisi (taulukko 4.4) ja toteutunut matriisi (taulukko 4.5). Kuusivaneritukin läpimitta on esitetty taulukon vasemmalla pystyrivillä millimetreinä ja sen vaihteluväli on 280 – 650 mm. Ylimmällä vaakarivillä taulukossa on kuusivaneritukin pituus senttimetreinä, siinä vaihteluvälinä on 330 – 610 cm. Taulukon sisäsolut näyttävät läpimitta-pituusyhdistelmien määrän.

Taulukko 4.6. Kuusivaneritukit, hintamatriisi.

VANERI	330	410	490	530	610
280	13,21	16,41	19,61	21,21	24,41
300	15,16	18,84	22,51	24,35	28,03
320	17,25	21,43	25,62	27,71	31,89
340	19,47	24,20	28,92	31,28	36,00
360	21,83	27,13	32,42	35,07	40,36
380	24,33	30,22	36,12	39,07	44,97
400	26,95	33,49	40,02	43,29	49,83
450	34,11	42,38	50,66	54,79	63,06
500	42,12	52,33	62,54	67,64	77,85
550	50,96	63,32	75,67	81,85	94,20
600	60,65	75,35	90,05	97,41	112,11
650	71,18	88,43	105,69	114,32	131,57

Kuusivaneritukkien hintamatriisi (taulukko 4.6) on laskettu samalla tavoin kuin sahatukkien vastaava soveltamalla kaavaa (4.1.). Kuusivaneritukin hinnaksi on oletettu 65 €/ m³.

4.4. Kuusen sahatukkiaineiston testitulokset

Aineistoon tehdyt testit on laskettu R-ohjelmiston avulla. Testeissä käytetyt R-koodit ovat liitteessä.

Taulukko 4.7. Kuusen sahatukkiaineiston testitulokset.

AI	1-C	Q	AP
0.347	0.051	0.459	0.183

Taulukosta 4.7 löytyvät kuusen sahatukkiaineiston testitulokset. Jakauma-asteen (AI) laskemiseen on käytetty kaavaa (3.1). Hintapainotettu jakauma-aste (AP) on laskettu käyttämällä kaavaa (3.3) ja suhteuttamalla se maksimiarvoonsa (kaava (3.4)). χ^2 -yhteensopivuustestiin (1-C) on käytetty kaavaa (3.6) ja se on suhteutettu kaavalla (3.7) ja vähentämällä saatu luku C ykkösestä. Laspeyresin määräindeksi (Q) on laskettu käyttämällä kaavaa (3.5).

4.5. Kuusivaneritukkiaineiston testitulokset

Aineistoon tehdyt testit on laskettu R-ohjelmiston avulla. Testeissä käytetyt R-koodit ovat liitteessä.

Taulukko 4.8. Kuusivaneritukkiaineiston testitulokset.

AI	1-C	Q	AP
0.287	0.074	0.459	0.157

Taulukossa 4.8. ovat kuusivaneritukkiaineiston testitulokset. Jakauma-asteen (AI) laskemiseen on käytetty kaavaa (3.1). Hintapainotettu jakauma-aste (AP) on laskettu käyttämällä kaavaa (3.3) ja suhteuttamalla se maksimiarvoonsa (kaava (3.4)). χ^2 -yhteensopivuustestiin (1-C) on käytetty kaavaa (3.6) ja se on suhteutettu kaavalla (3.7) ja vähentämällä saatu arvo C ykkösestä. Laspeyresin määräindeksi (Q) on laskettu käyttämällä kaavaa (3.5).

χ^2 -yhteensopivuustestissä sekä Laspeyresin määräindeksissä on käytetty suhteutettua toteutunutta matriisia eli jokainen on solu on jaettu tukkien kokonaismäärällä ja kerrottu tuhannella. Näin suhteutettu toteutunut matriisi on vertailukelpoinen tavoitematriisiin kanssa. Jakauma-astetta tai sen hintapainotettua versiota käytettäessä suhteutusta ei tarvitse tehdä, koska kaava itsessään sisältää sen. Näin ollen tulokset ovat samoja käytettiin sitten suhteutettua tai suhteuttamatonta toteutunutta matriisia. Suhteutetut toteutuneet matriisit esitetään luvussa 4.8 (taulukot 4.11 ja 4.12).

4.6. Jakauma-asteen odotusarvot

Jakauma-asteen odotusarvot $E(AI)$ saavat arvot 0.974 (kuusen sahatukki) ja 0.962 (kuusivaneritukki). Odotusarvot on laskettu soveltaen kaavaa (3.22) ja niiden laskemiseen on käytetty R-ohjelmaa. Ohjelmassa käytetty koodi on liitteessä.

Odotusarvot $E(AI)$ ovat korkeita, koska tukkien määrä (N) toteutuneessa matriisissa on melko suuri; kuusen sahatukilla $N = 19\,997$ ja kuusivaneritukilla $N = 4\,950$. Kun tukkien määrä kasvaa, lähenee $E(AI)$ lukua yksi (Koskela et al. 2006).

4.7. Yhteensopivuusmittojen tulokset leimikoille erikseen

Kuten aineiston esittelystä (kappale 4.1) käy ilmi, on aineistossa käytetty 15 eri leimikkoa. Laskin yhteensopivuusmittojen tunnusluvut myös jokaiselle leimikolle erikseen sekä kuusen sahatukin että kuusivaneritukin osalta.

Taulukko 4.9. Kuusen sahatukkien tulokset leimikoille.

Leimikko	AI	1-C	Q	AP	N
1	0.362	0.051	0.473	0.192	1560
2	0.344	0.043	0.450	0.179	2287
3	0.306	0.043	0.430	0.156	751
4	0.359	0.057	0.478	0.192	548
5	0.357	0.054	0.483	0.194	668
6	0.347	0.052	0.459	0.183	2905
7	0.349	0.046	0.460	0.186	2930
8	0.330	0.045	0.444	0.173	681
9	0.359	0.054	0.453	0.187	2193
10	0.326	0.051	0.463	0.171	610
11	0.356	0.049	0.466	0.188	1146
12	0.351	0.049	0.451	0.186	1309
13	0.372	0.050	0.459	0.196	521
14	0.271	0.037	0.475	0.145	456
15	0.276	0.040	0.463	0.146	1432

Taulukko 4.10. Kuusivaneritukkien tulokset leimikoille erikseen.

Leimikko	AI	1-C	Q	AP	N
1	0.174	0.046	0.405	0.079	148
2	0.179	0.039	0.406	0.082	118
3	0.230	0.051	0.428	0.116	83
4	0.294	0.084	0.470	0.161	181
5	0.284	0.079	0.460	0.151	350
6	0.219	0.059	0.430	0.110	492
7	0.257	0.062	0.443	0.139	514
8	0.249	0.077	0.455	0.133	159
9	0.181	0.047	0.411	0.089	362
10	0.216	0.053	0.425	0.108	183
11	0.262	0.067	0.454	0.140	367
12	0.273	0.065	0.468	0.146	422
13	0.441	0.121	0.557	0.290	457
14	0.266	0.074	0.459	0.137	222
15	0.320	0.080	0.472	0.176	892

Taulukoissa 4.9 ja 4.10 on esitetty yhteensopivuusmittojen tulokset jokaiselle leimikolle erikseen. Jokaisen testin suurin yhteensopivuusarvo 15 leimikon joukossa on lihavoitu ja vastaavasti pienin arvo on kursivoitu. Taulukoissa N kertoo kyseessä olevasta leimikosta kaadettujen puiden kokonaismäärän.

4.8. Yhteensopivuusmittojen tulosten tulkintaa

Toteutunutta matriisia sekä tavoitematriisia vertaillen molemmissa tapauksissa huomataan, ettei yhteensopivuus voi olla kovin suuri, koska kummankaan tapauksessa ylemmät läpimittatavoitteet eivät ole toteutuneet lainkaan (ko. solujen arvot ovat nolliä). Lisäksi alimmat läpimittaluokat ovat "ylitoteutuneita" eli niitä saadaan tavoitteisiin nähden selkeästi liikaa.

Taulukko 4.11. Sahatukit, suhteutettu toteutunut matriisi.

SAHATUKKI	370	430	460	490	520	550	580	610	Yhteensä
160	56	33	37	37	38	35	41	123	400
170	4	6	5	14	4	9	5	6	53
180	53	5	2	7	2	4	1	3	78
200	94	26	10	12	6	11	2	4	165
220	0	111	20	18	9	15	5	4	182
240	0	57	12	7	6	8	1	2	93
260	0	18	2	2	5	3	0	0	29
280	0	0	0	0	0	0	0	0	0
300	0	0	0	0	0	0	0	0	0
320	0	0	0	0	0	0	0	0	0
340	0	0	0	0	0	0	0	0	0
360	0	0	0	0	0	0	0	0	0
380	0	0	0	0	0	0	0	0	0
400	0	0	0	0	0	0	0	0	0
Yhteensä	207	256	87	97	69	84	55	143	1000

Taulukko 4.12. Kuusivaneritukit, suhteutettu toteutunut matriisi.

kuusivaneri	330	410	490	530	610	Yhteensä
280	115	93	105	93	130	535
300	72	11	8	2	1	93
320	88	27	19	8	2	143
340	54	15	16	5	4	93
360	34	6	7	4	2	53
380	29	7	6	2	1	45
400	21	3	5	1	1	31
450	3	0	2	0	0	6
500	0	0	1	0	0	1
550	0	0	0	0	0	0
600	0	0	0	0	0	0
650	0	0	0	0	0	0
Yhteensä	415	162	168	115	141	1000

Tilastollinen χ^2 -yhteensopivuustesti rankaisee tästä ankarimmin antaen hyvin pienet yhteensopivuusarvot; 0.051 ja 0.074. Tämä testi ei ota hintamatriisia huomioon ollenkaan. Hintapainotettu jakauma-aste antaa pienemmät arvot kuin jakauma-aste, koska arvokkaimpia (tilavuudeltaan suurempia) tukkeja ei kummassakaan luokassa saada tai niitä saadaan hyvin vähän (toteutuneessa matriisissa nämä solut ovat arvoltaan nollia tai lähellä nollaa).

Suurimmat yhteensopivuusarvot antaa yleisesti kansantaloustieteessä käytetty Laspeyresin määräindeksi, vaikka sekin ottaa hintamatriisin

huomioon. Tosin pitää muistaa, että Laspeyresin indeksin tunnusluvun vaihteluväli on erilainen kuin muilla testeillä eli se voi ylittää luvun yksi.

Näin pienet yhteensopivuusarvot olivat siis odotettavissa, koska tavoitematriisin suurimpien luokkien tavoitearvot eivät ole täyttyneet ollenkaan. Yksi syy tähän voisi olla leimikon rakenne. Esimerkiksi sellainen leimikko, jossa ei ole lainkaan korkeampia läpimitta-asteita täyttyviä puita, voisi johtaa tällaiseen toteutuneeseen matriisiin.

Kun tarkastellaan aineiston kaikkia leimikkoja erikseen (taulukot 4.9 ja 4.10), huomataan ettei puiden määrällä (N) ole ollut juurikaan vaikutusta mittojen arvoihin. Taulukossa 4.13 esitellään käytettyjen yhteensopivuusmittojen tunnuslukujen vaihteluvälit eli pienin ja suurin arvo, jonka ko. tunnusluku on saanut näiden 15 leimikon joukossa. Vaihteluvälit esitellään kuusen sahatukille ja kuusivaneritukille erikseen.

Taulukko 4.13. Yhteensopivuusmittojen tunnusarvojen vaihteluvälit.

	Sahatukki	Kuusivaneritukki
AI	0.271, 0.372	0.174, 0.441
1-C	0.037, 0.057	0.039, 0.121
Q	0.430, 0.483	0.405, 0.557
AP	0.145, 0.196	0.079, 0.290

Suurin vaihteluväli eli pienimmän ja suurimman tunnusarvon välinen erotus oli jakauma-asteen (AI) tunnusluvuissa (sahatukki 0.101, kuusivaneritukki 0.267). Hintapainotetun jakauma-asteen (AP) tunnusarvojen erotus oli jonkin verran pienempi molemmissa tapauksissa (0.051, 0.211). Pienin vaihteluväli oli X^2 -yhteensopivuustestin tunnusluvulla 1-C (sahatukki 0.020, kuusivaneritukki 0.082).

5. Simulointi

Tässä luvussa vertaillaan jakauma-asteen, hintapainotetun jakauma-asteen, χ^2 -yhteensopivuustestin sekä Laspeyresin määräindeksin tunnuslukuja sekä sitä, miten ne käyttäytyvät, kun tehdään normaalijakaumaan pohjautuva simulointi ja kasvatetaan normaali-jakauman varianssia.

Jakauma-asteen laskemiseen käytetään yhtälöä (3.1) ja χ^2 -yhteensopivuustestin tunnuslukuna käytetään ykkösestä vähennettyä kontingenssikerrointa (1-C) (ks. luku 3.5). Hintapainotetun jakauma-asteen tunnusluku on AP/AP_{MAX} eli yhtälö (3.3) suhteutettuna yhtälöön (3.4). Laspeyresin määräindeksin tunnusluku on Q_L (yhtälö 3.5).

Tarkoituksena on normaalijakaumaa apuna käyttäen simuloida 1 000 kappaletta erilaisia toteutuneita matriiseja ja laskea jokaisesta mainittujen yhteensopivuusmittojen tunnusluvut. Näistä 1 000 simuloinnista lasketaan sitten tunnuslukujen keskiarvot ja luottamusvälit. Simuloinnissa käytetään normaalijakaumaa apuna, joka esitellään seuraavaksi.

5.1. Normaalijakauma

Normaalijakauma on jatkuvista jakaumista tunnetuin ja käytetyin. Sen varhaisimpia kehittäjiä olivat 1700-luvulla eläneet ranskalaiset matemaatikot Abraham de Moivre ja Pierre Simon Laplace. Saksalainen matemaatikko Carl Friedrich Gauss sovelsi normaalijakaumaa tähtitieteellisten mittausvirheiden käsittelyyn. (Huuhtanen, Kallinen 1998.)

Satunnaismuuttujan X jakauma on normaali parametrein μ , $-\infty < \mu < \infty$, ja σ^2 , $\sigma^2 > 0$, $X \sim N(\mu, \sigma^2)$, jos X :n tiheysfunktio on

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2},$$

jossa μ on satunnaismuuttujan X odotusarvo ja σ^2 sen varianssi (Huuhtanen, Kallinen 1998).

5.2. Simulointialgoritmi

Simuloinnissa käytetään pohjana kuusivaneritukkien tavoitematriisia (taulukko 4.1). Jokaiselle läpimitta-pituusluokalle generoidaan arvo (toteutunut tukkien lukumäärä ko. luokalle) normaalijakaumasta, jossa odotusarvona on ko. läpimitta-pituusluokan tavoitematriisin arvo ja varianssina tietty vakio. Tällöin toteutuneen matriisin solujen summan (tukkien määrän) odotusarvo on tietysti sama kuin tavoitematriisin (tässä tapauksessa 1 000), mutta käytännössä se vaihtelee hieman riippuen määrittelystä varianssin arvosta.

Simuloinnit on tehty R-ohjelmalla. Koodissa on otettu huomioon tavoitematriisissa luokan (solun) kohdalla arvo ollessa nolla, myös toteutuneeseen matriisiin jätetään silloin arvo nolla. Lisäksi varmistetaan, ettei toteutuneeseen matriisiin tule negatiivisia arvoja (eli tukkien määrä ei missään solussa voi olla nollaa pienempi). Jokaisen generoidun toteutuneen matriisin arvot myös pyöristetään lähimpään kokonaislukuun. Tässä simuloinnissa tehdään 1 000 toistoa jokaisella käytetyllä varianssin arvolla. Käytetty R-koodi on liitteestä.

5.3. Simuloinnin tulokset

Simuloinnin tulokset on esitetty taulukoissa 5.1 – 5.4. Jokaisessa kokeessa (eri varianssien tapauksissa) tehtiin siis 1 000 toistoa ja jokaisesta simuloinnista laskettiin kaikkien neljän tunnusluvun (AI, AP, Q ja 1-C) alaraja, keskiarvo sekä yläraja. Alaraja ja yläraja on laskettu

järjestämällä toistoista saadut tunnusluvut (1 000 kpl) pienimmästä suurimpaan ja valitsemalla 25. havainto (alaraja) sekä 975. havainto (yläraja). Näin ollen alarajan ja ylärajan väliin jää 95 % havainnoista (95 % luottamusväli).

Taulukko 5.1. Jakauma-asteen simulointitulokset

Varianssi	AI alaraja	AI ka	AI yläraja
1	0,973	0,978	0,983
3	0,918	0,934	0,947
5	0,868	0,891	0,912
9	0,783	0,821	0,855

Taulukko 5.2. Hintapainotetun jakauma-asteen simulointitulokset

Varianssi	AP alaraja	AP ka	AP yläraja
1	0,972	0,979	0,986
3	0,917	0,937	0,954
5	0,864	0,897	0,924
9	0,776	0,830	0,873

Taulukko 5.3. χ^2 -yhteensopivuustestin simulointitulokset

Varianssi	1-C alaraja	1-C ka	1-C yläraja
1	0,915	0,930	0,943
3	0,758	0,800	0,835
5	0,635	0,686	0,737
9	0,464	0,528	0,592

Taulukko 5.4. Laspeyresin määräindeksin simulointitulokset

Varianssi	Q alaraja	Q ka	Q yläraja
1	0,985	1,000	1,017
3	0,953	1,000	1,049
5	0,925	1,000	1,077
9	0,904	1,026	1,158

5.4. Simulointituloksien tulkintaa

Jakauma-asteen odotusarvo tässä tapauksessa (toteutuneen jakauman tukkien lukumäärä $\approx 1\,000$) kuusivaneritukille on 0,914. Odotusarvo ylittyy jakauma-asteen ja hintapainotetun jakauma-asteen tunnusluvuissa, kun varianssin arvo on yksi ja kolme. Varianssin kasvaessa tunnusluku odotetusti pienenee sekä jakauman luottamusväli kasvaa.

Hintapainotus ei näytä vaikuttavan näiden kahden testisuureen välillä vaan simulointitulokset ovat hyvinkin samansuuntaisia (katso taulukot 5.1 ja 5.2). Tämä johtuu suureksi osaksi siitä, että tämä simulointi olettaa kaikille pituus-läpimittaluokille yhtä suuret varianssit ja odotusarvo on juuri se mikä tavoitematriisissa on annettu. Normaalisissa tilanteissa leimikon rakenne saattaisi esimerkiksi olla sellainen, ettei siellä välttämättä olisi suurempiin läpimitta-pituusluokkiin sopivia puita niin paljon. Tällaisessa tapauksessa hintapainotetun jakauma-asteen tunnusluvut jäisivät todennäköisesti pienemmiksi kuin jakauma-asteen. Lisäksi hintamatriisi on tässä tapauksessa yksinkertaistettu, koska siinä tilavuus vaikuttaa suoraan tukin arvoon; mitä suurempi tilavuus, sitä arvokkaampi tukki.

Verrattuna muihin tunnuslukuihin, χ^2 -yhteensopivuustestin (taulukko 5.3) tunnusluvut 1-C laskevat nopeimmin varianssin kasvaessa. Lisäksi tämän tunnusluvun luottamusväli laajenee eniten varianssin kasvaessa verrattuna muihin tunnuslukuihin.

Laspeyresin määräindeksi Q (taulukko 5.4) antaa varianssien yksi, kolme ja viisi arvoilla keskiarvoksi luvun yksi (varianssin kasvaessa yhdeksään 1,026). Varianssin kasvaessa luottamusväli laajenee. Laspeyresin indeksi on näistä testeistä ainoa, jossa testin tunnusarvo voi ylittää luvun yksi, mikä vaikeuttaa tämän tunnusluvun vertailua χ^2 -yhteensopivuustestin sekä jakauma-asteiden tunnuslukuihin. Tulosten perusteella voidaan myös todeta, että mitä lähempänä luku yksi

Laspeyresin määräindeksi on, sitä paremmin toteutunut jakauma vastaa tavoitejakaumaa.

Varianssin kasvaessa (katso taulukot 5.1 – 5.4) yleinen trendi on, että luottamusväli laajenee sekä tunnusluku pienenee. Näin tapahtuu selkeästi jakauma-asteen, hintapainotetun jakauma-asteen sekä χ^2 -yhteensopivuustestin tapauksessa. Ainoan poikkeuksen muodostaa Laspeyresin määräindeksi; sen tapauksessa luottamusväli kyllä laajenee, mutta sen arvot pyörivät selkeästi luvun yksi ympärillä.

Jakauma-aste näyttäisi olevan käytetyistä yhteensopivuusmitoista selkein, koska se on helpoiten tulkittavissa oleva. Sen tunnusluku kertoo kuinka monta prosenttia tukeista on osunut oikeisiin läpimittapituusluokkiin (Nummi, Sinha, Koskela 2006). Jakauma-asteen tunnuslukuihin verrattuna χ^2 -yhteensopivuustesti näyttää antavan pienempiä arvoja, jotka voidaan mahdollisesti tulkita liian ankarasti ts. yhteensopivuus voi olla hyväksyttävä paljon pienemmillä tunnusluvun 1-C arvoilla kuin jakauma-asteen tapauksessa. Laspeyresin määräindeksin huonona puolena on sen jossain määrin heikompi tulkittavuus muihin yhteensopivuusmitoihin verrattuna.

Lähteet

Huhtanen, P. & Kallinen, A. (1998). Matemaattinen tilastotiede. Tampereen yliopisto.

Inman, H.F. & Bradley, E.L. (1989). The Overlapping Coefficient as a Measure of Agreement Between Probability Distributions and Point Estimation of the Overlap of Two Normal Densities. *Commun. in Statist. – Theory Meth.* 18, s. 3851-3874

Kauhanen, J. (1997): Metsäteollisuus Suomessa. Metsät Luonnonvarana - Turun Yliopiston Maantieteen Laitoksen Monisteita N:o 15, s. 9-14

Kivinen, V. –P. (2004): Katkonnanohjaus koneellisessa hakkuussa. Esitys Mikkelin ammattikorkeakoulussa 16.12.2004. Moniste.

Kivinen, V.-P., Uusitalo, J. & Nummi, T. (2005). Comparison of Four Measures Designed for Assessing the Fit between the Demand and Output Distribution of Logs. *Canadian Journal of Forest Research* Vol. 35, s. 693-702

Koskela, L., Sinha, B. & Nummi, T. (2006). Some Aspects of the Sampling Distribution of the Apportionment Index and Related Inference. Julkaisematon käsikirjoitus

MAOL-taulukot (1995). Kustannusosakeyhtiö Otava

Metsänhoitoyhdistys Keskipohja.

<http://www.mhy.fi/keskipohja/puumarkkinat/fi_FI/index/>. Internet-lähde

Metsäteollisuus ry (2000). Avain Suomen metsäteollisuuteen. Libris Oy, Helsinki

Metsäteollisuus ry

<<http://www.forestindustries.fi/index.html>>. Internet-lähde

Nummi, T., Sinha, B.K. & Koskela, L. (2005). Statistical Properties of the Apportionment Degree and Alternative Measures in Bucking Outcome. *Revista Investigación Operacional*, Vol. 26, No. 3.

Nummi, T., Sinha, B.K. & Koskela, L. (2006). Some Aspects of the Sampling Distribution of the Apportionment Index and Related Inference. *Konferenssiesitys 2006, Havanna, Kuuba.*

Uusitalo, J. & Kivinen, V.-P. (2004) Introduction to tree bucking control (Katkonnanhjauksen perusteet). Yhteenveto luennoista "Tree bucking control" -kurssilta Joensuun yliopistossa 22.11. -1.12.2004.

Weitzman, M.S. (1970). Measures of Overlap of Income Distributions of White and Negro Families in the United States. Technical Paper no. 22. Washington: U.S. Department of Commerce, Bureau of the Census.


```

target.t<-sahatukki.demand.matrix
output.f<-sahatukki.output.matrix
price.p<-sahatukki.price.matrix
output.f.suht<-sahatukki.output.matrix.suht

#jakauma-aste lasketaan soveltamalla kaavaa (3.1)

A=1-0.5*sum(abs(output.f/sum(output.f)-target.t/sum(target.t)))

A
#[1] 0.3465786

# $\chi^2$ -yhteensopivuustesti soveltamalla kaavoja (3.6) ja (3.7)

target.t<-target.t+0.0001 #target-matriisissa nollat pois, ettei tule nolalla jakoa

chi.sq<-sum((output.f.suht-target.t)^2/target.t)

N<-sum(output.f.suht)

C<-sqrt(chi.sq/(chi.sq+N)) # N=sum(output.f) = 19 997

C

# 0.949301

1-C

#0.05069903

#hintapainotettu jakauma-aste soveltaen kaavoja (3.3) ja (3.4)

omega.ij<-(price.p/(sum(price.p)))

x.omega<-sum(output.f/sum(output.f)*omega.ij)

theta.omega<-sum(target.t/sum(target.t)*omega.ij)

```

```
A.p<-(x.omega+theta.omega)/2-0.5*sum(omega.ij*abs(output.f/sum(output.f)-
target.t/sum(target.t)))
```

```
A.p.max<-sum(target.t/sum(target.t)*omega.ij)
```

```
A.p
```

```
# 0.001539065
```

```
A.p.max
```

```
# 0.008412918
```

```
A.p/A.p.max
```

```
# 0.1829407
```

```
#Laspeyresin määräindeksi soveltaen kaavaa (3.5)
```

```
Q.L<-(sum(price.p*output.f.suht))/(sum(price.p*target.t))
```

```
Q.L
```

```
# 0.4591137
```

```
#KUUSIVANERITUKKIEN TESTIT
```

```
#tässä syötetään koneelle sahatukkien tavoitejakaumat, toteutuneet jakaumat sekä hintamatriisi.
```

```
vaneri.price.matrix<-
```

```
matrix(c(13.21,16.41,19.61,21.21,24.41,15.16,18.84,22.51,24.35,28.03,17.25,21.43,25.62,27.71,3
1.89,19.47,24.20,28.92,31.28,36.00,21.83,27.13,32.42,35.07,40.36,24.33,30.22,36.12,39.07,44.97,
26.95,33.49,40.02,43.29,49.83,34.11,42.38,50.66,54.79,63.06,42.12,52.33,62.54,67.64,77.85,50.9
6,63.32,75.67,81.85,94.20,60.65,75.35,90.05,97.41,112.11,71.18,88.43,105.69,114.32,131.57),nco
l=12,nrow=5)
```

```
vaneri.price.matrix<-t(vaneri.price.matrix)
```

```
vaneri.output.matrix<-
```

```
matrix(c(568,458,518,461,642,354,52,40,9,7,436,132,93,39,9,265,74,78,26,19,167,31,35,18,9,142
,36,28,11,4,102,15,26,7,4,16,2,9,0,2,2,1,3,0,0,0,0,0,0,0,0,0,0,0,0,0,0),ncol=12,nrow=5)
```

```

vaneri.output.matrix<-t(vaneri.output.matrix)

vaneri.demand.matrix<-
matrix(c(10,10,40,10,10,10,10,40,10,10,10,10,40,10,10,10,10,40,10,10,10,10,60,1
0,10,10,10,60,10,10,10,10,60,10,10,10,10,60,10,10,10,10,60,10,10,0,0,0,0),ncol=
12,nrow=5)

vaneri.demand.matrix<-t(vaneri.demand.matrix)

vaneri.output.matrix.suht<-
matrix(c(115,93,105,93,130,72,11,8,2,1,88,27,19,8,2,54,15,16,5,4,34,6,7,4,2,29,7,6,2,1,21,3,5,1,1,
3,0,2,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),ncol=12,nrow=5)
vaneri.output.matrix.suht<-t(vaneri.output.matrix.suht)

target.t<-vaneri.demand.matrix
output.f<-vaneri.output.matrix
price.p<-vaneri.price.matrix
output.f.suht<-vaneri.output.matrix.suht

#Jakauma-aste lasketaan soveltamalla kaavaa (3.1)

A<-1-0.5*sum(abs(output.f/sum(output.f)-target.t/sum(target.t)))

A

# 0.2869697

# $\chi^2$ -yhteensopivuustesti soveltamalla kaavoja (3.6) ja (3.7)

target.t<-target.t+0.0001 #target-matriisissa ei voi olla nollia, ettei tule nolllalla jakoa

chi.sq<-sum((output.f.suht-target.t)^2/target.t)

N<-sum(output.f.suht)

C<-sqrt(chi.sq/(chi.sq+N))

```

C

0.9255257

1-C

0.07447432

#hintapainotettu jakauma-aste soveltaen kaavoja (3.3) ja (3.4)

$\omega_{ij} \leftarrow (\text{price}_p / (\sum(\text{price}_p)))$

$x.\omega \leftarrow \sum(\text{output}_f / \sum(\text{output}_f) * \omega_{ij})$

$\theta.\omega \leftarrow \sum(\text{target}_t / \sum(\text{target}_t) * \omega_{ij})$

$A.p \leftarrow (x.\omega + \theta.\omega) / 2 - 0.5 * \sum(\omega_{ij} * \text{abs}(\text{output}_f / \sum(\text{output}_f) -$

$\text{target}_t / \sum(\text{target}_t)))$

$A.p.\text{max} \leftarrow \sum(\text{target}_t / \sum(\text{target}_t) * \omega_{ij})$

A.p

0.002483129

A.p.max

0.01578843

A.p/A.p.max

0.1572753

#Laspeyresin määräindeksi soveltaen kaavaa (3.5)

$Q.L \leftarrow (\sum(\text{price}_p * \text{output}_f) / (\sum(\text{price}_p * \text{target}_t)))$

Q.L

0.4592971

```
#JAKAUMA-ASTEEN ODOTUSARVOT
```

```
# Lasketaan E(AI) sahatukille
```

```
target<-sahatukki.demand.matrix
```

```
target.tahti<-target/(sum(target))
```

```
N<-19997
```

```
E.AI<-1-sqrt(1/(2*N*pi))*sum(sqrt(target.tahti*(1-target.tahti)))
```

```
E. AI
```

```
# [1] 0.9735704
```

```
# Lasketaan E(AI) vanerille
```

```
target<-vaneri.demand.matrix
```

```
target.tahti<-target/(sum(target))
```

```
N<-4950
```

```
E.AI<-1-sqrt(1/(2*N*pi))*sum(sqrt(target.tahti*(1-target.tahti)))
```

```
E.AI
```

```
# [1] 0.9615401
```

```

# SIMULOINTI
target<-vaneri.demand.matrix
target<-target+0.0001    #ei nollia target-matriisissa
price<-vaneri.price.matrix
omega.ij<-(price/(sum(price)))

toistot<-1000 #toistojen lkm
V<-9        #variassi

jakauma.aste<-rep(0,toistot)
chi<-rep(0,toistot)
lasp<-rep(0,toistot)
jakauma.price<-rep(0,toistot)

for (k in 1:toistot) #luuppi alkaa

{
output<-matrix(c(rep(0,60)),nrow=12,ncol=5)

for (i in 1:nrow(target)){
for (j in 1:ncol(target)){

if (target[i,j]==0.0001)
    output[i,j]<-0
else
    output[i,j]<-rnorm(1, target[i,j], V)
}
}

for (i in 1:nrow(output)){
for (j in 1:ncol(output)){

if (output[i,j]<0)
    output[i,j]<-0
else
    output[i,j]<-output[i,j]
}
}
}

```



```
jakauma.aste[k]<-(1-0.5*sum(abs(round(output)/sum(round(output))-target/sum(target))))
```

```
chi.sq<-sum((round(output)-target)^2/target)
```

```
chi[k]<-(1-sqrt(chi.sq/(chi.sq+sum(round(output)))))
```

```
lasp[k]<-(sum(price*round(output)))/(sum(price*target))
```

```
x.omega<-sum(round(output)/sum(output)*omega.ij)
```

```
theta.omega<-sum(target/sum(target)*omega.ij)
```

```
A.p<-(x.omega+theta.omega)/2-0.5*sum(omega.ij*abs(round(output)/sum(round(output))-  
target/sum(target))
```

```
A.p.max<-sum(target/sum(target)*omega.ij)
```

```
jakauma.price[k]<-A.p/A.p.max
```

```
}
```

```
c(jakauma.aste[order(jakauma.aste)][25], mean(jakauma.aste),  
jakauma.aste[order(jakauma.aste)][975])
```

```
c(jakauma.price[order(jakauma.price)][25], mean(jakauma.price),  
jakauma.price[order(jakauma.price)][975])
```

```
c(chi[order(chi)][25], mean(chi), chi[order(chi)][975])
```

```
c(lasp[order(lasp)][25], mean(lasp), lasp[order(lasp)][975])
```