

ESIPUHE

Seuraavassa julkaistaan kaksi ETLA:n tutkijoiden metodologista työtä. Valtiot.kand., luonnont.kand. Yrjö O. Vartian tutkimus käsittelee erään laajan jakaumaperheen ns. betajakauman parametrien estimointia maximum likelihood menetelmällä. Tutkimus liittyy päätösvaihtoehtojen paremmuusjärjestyksen määrittämisen yleiseen ongelmaan.

Valtiot.lis., dipl.ins. Pentti L.I. Vartian tutkimus käsittelee tieteellisten teorioiden ns. identifiointikysymystä. Identifiointiongelma aiheutuu siitä tosiasiasta, että havainnot voidaan usein selittää monella eri tavalla. Taloudellisia malleja rakennettaessa identifiointi vaikuttaa sekä mallin toiminnan tulkitaan että sen parametrien estimointiin.

Tieteellisen antinsa ohella tämänlaatuinen perustutkimus muodostaa edellytyksen pitkäjänteiselle tutkimustoiminnalle. On myös selvää, että metodisen perustutkimuksen antama valmius on usein tarpeen päiväkohtaisessakin selvitystyössä.

Konekirjoituksen on suorittanut ja kuvat piirtänyt Ann-Christine Ekebohm.

Helsingissä lokakuussa 1973

Kaarlo Larna

YRJÖ OLAVI VARTIA

BETAJAKAUMAN PARAMETRIEN ESTIMOINTI
MAXIMUM LIKELIHOOD-MENETELMÄLLÄ

ESTIMATION OF THE PARAMETERS OF THE BETA
DISTRIBUTION BY MAXIMUM LIKELIHOOD METHOD

Helsinki 1973

SISÄLLYSLUETTELO

	sivu
1. JOHDANTO	7
2. BETAJAKAUMAN OMINAISUUDET	10
2.1. Yleistä	10
2.2. Tiheysfunktion käyttäytyminen	13
2.3. Tunnusluvut	18
2.4. Maximum likelihood-yhtälöt: Kirjoitustapoja ja tulkintoja	23
2.5. Betajakauman ja eräiden muiden jakaumien välisiä yhteyksiä	31
3. ERÄS SOVELLUTUSALUE: PÄÄTÖSVAIHTOEHTOJEN PAREMMUUSJÄRJESTYKSEN MÄÄRÄÄMINEN	35
4. BETAJAKAUMAN ML-YHTÄLÖIDEN RATKAISEMINEN ELI BIJEKTION $(a,b) \rightarrow (f,g)$ KÄÄNTEISFUNKTION APPROKSIMOINTI	52
4.1. Yleistä	52
4.2. Normaalijakauman muotoisten betajakaumien tapaus: $a > 2$ ja $b > 2$ sekä $a/b \approx 1$	61

	sivu
4.3. U-muotoisten betajakaumien tapaus: $a < 1$ ja $b < 1$	73
4.4. Gammajakauman muotoisten betajakaumien tapaus: $a/b \approx 0$ tai $b/a \approx 0$	79
 KIRJALLISUUSVIITTEET	 86
 LIITTEET	

LIITTEET

	sivu
1. Psii-funktion ja psii-funktion käänteisfunktion approksimointi	88
2. M1-yhtälöiden ratkaisu symmetrisessä tapauksessa $a = b$.	102
3. M1-yhtälöiden iteratiivinen ratkaisu ja siihen liittyvä tietokoneohjelma	109
4. Likelihood yhtälöiden ratkaisuja $\theta = 0.1, 0.2, \dots, 0.9$ ja $p = 0.1, 0.2, \dots, 0.5$ vastaavat havaintoarvot u ja v sekä f ja g	113
5. Joitakin tärkeitä vakioita	114
6. Arvioiden $1/x \approx 2d$ ja $p \approx \frac{f+d}{1+d}$ tarkkuudet	115
7. Arvioiden $1/x \approx \sqrt{uv}$ ja $p \approx 1/(1+\sqrt{\frac{u}{v}})$ tarkkuudet	116
8. Arvioiden z_0 ja z_2 tarkkuudet	117
9. Arvioiden a_2 ja b_2 tarkkuudet	118
10. Arvioiden a_3 ja b_3 tarkkuudet	119
11. Arvioiden a_4 ja b_4 tarkkuudet	120

1. JOHDANTO

Jakaumamallin sovittaminen havaintoaineistoon on tärkeä apuväline useilla tilastotieteen sovellutusaloilla kuten vakuutustoiminnassa, laaduntarkkailussa ja luotettavuuslaskelmissa, joissa päätökset perustuvat usein laajoihin empiirisiin aineistoihin. Tässä työssä tarkastellaan betajakaumaa ja sen parametrien estimointia maximum likelihood-menetelmällä. Tuloksilla on suoraan käyttöä järjestetyn otoksen havaintojen avulla formuloitujen jakaumamallien analysoinnissa ja tältä pohjalta voidaan kehittää mielenkiintoinen ja nähdäkseni huomattavan yleinen kehikko päätösvaihtoehtojen paremmuusjärjestyksen määrittämiseksi.

Betajakauma sisältää raja- tai erikoistapauksenaan mm. normaali-jakauman, gammajakauman ja tasajakauman, joten betajakaumalle johdettuja tuloksia voidaan soveltaa likimääräisinä myös näille jakaumille. Toisaalta betajakauman avulla voidaan tutkia, kuinka tilastolliset menetelmät yleistyvät siirryttäessä tavanomaisista jakaumamalleista mutkikkaampaan malliin. Päätösvaihtoehtojen paremmuusjärjestyksen määrittäminen tarjoaa tästä erään konkreettisen esimerkin.

Tutkimuksen nimi ei ehkä anna oikeata kuvaa käsiteltävästä ongelmasta. Koska betajakauman tapauksessa ml-yhtälöt merkitsevät

teoreettiset ja empiiriset geometriset keskiarvot yhtä suuriksi, on matemaattisena ongelmana itse asiassa selvittää erään laajan jakaumaperheen geometrinen keskiarvojen $f = \text{Geom}(x)$ ja $g = \text{Geom}(1-x)$ ja muiden parametrien väliset riippuvuudet likimääräisessä, mutta eksplisiittisessä muodossa. Numeeristen ratkaisumenetelmien avulla voidaan nämä riippuvuudet aina tapaus tapaukselta laskea¹⁾, mutta näin saadut numeeriset vastaavuudet ovat teoreettisissa tarkasteluissa ja riippuvuuksien ymmärtämisen kannalta tavallisesti käyttökeltvottomia.

Laajennettaessa tulosten käyttöä betajakauman rajajakaumiin, havaitaan betajakauman ml-yhtälöiden täsmällisen ratkaisun turhuus. Jokainen malli on käytännössä vain likimääräinen, joten malliin liittyviä ongelmia ei yksittäisessä sovellutuksessa tarvitse ratkaista täydellä tarkkuudella. Liiallisella täsmällisyydellä operoitaessa menetetään usein tulosten yleisemmät tulokset ja sovellutukset. Tutkimuskohteena eivät useinkaan ole yksityisen mallin tarkat ominaisuudet, vaan vaihtoehtoisten mallien välisten erojen ja yhtäläisyyksien vertailu ja tähän perustuva käyttökelpoisimman ratkaisutavan valinta.

Työ on laajennus aihetta käsittelevästä tilastotieteen laudaturtyöstä, joka syntyi työskennellessäni vuosina 1970-71 Posti- ja lennätinlaitoksen Liiketaloudellisessa tutkimuslaitoksessa. Eri-tyisesti luku 4.4. on kirjoitettu uudestaan uusien tutkimustulosten pohjalta. Samoin kolmas, päätösvaihtoehtojen paremmuusjärjestystä käsittelevä luku on lisäys aikaisempaan. Haluan kiittää prof. Leo Törnqvistiä ohjauksesta, joka on avannut uusia ja

1) Näin on tehty artikkelissa Gnadesikon, Pinkham, Hughes (1).
Liitteessä 3 on esitetty toinen, tehokas ratkaisualgoritmi.

hyödyllisiä näköaloja matemaattisten kysymysten ratkaisemiseen sekä VTL, dipl.ins. Pentti L.I. Vartiaa useista keskusteluista, jotka ovat selventäneet ja konkretisoineet työssä tarkasteltuja, usein abstrakteilta tuntuvia kysymyksiä.

2. BETAJAKAUMAN OMINAISUUDET

2.1. YLEISTÄ

Pearsonin jakaumasysteemin I tyyppi sisältää jakaumat, joiden parametreistä $p > 0$ ja $q > 0$ riippuva tiheysfunktio on muotoa

$$(2.1.1) \quad \frac{\Gamma(p+q)}{(a_1+a_2)^{p+q-1} \Gamma(p) \Gamma(q)} (a_1+x)^{p-1} (a_2-x)^{q-1} \phi(-a_1 < x < a_2) \text{ ja}$$

$$(2.1.2) \quad \Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt = \text{Eulerin gammafunktio}$$

$$(2.1.3) \quad \phi(A) = \text{joukon } A \text{ indikaattorifunktio}$$

Tavallisesti havaintojen variaatioväli $(-a_1, a_2)$ muunnetaan lineaarimuunnoksella väliksi $(0,1)$.

Seuraavassa sanotaankin satunnaismuuttujan x noudattavan betajakaumaa parametrein (a,b) , jos sen tiheysfunktio on muotoa:

$$(2.1.4) \quad \beta(x|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} \phi(0 < x < 1),$$

jossa $a > 0$, $b > 0$.

Tällöin merkitään $x \sim \text{Beta}(a,b)$.

Seuraavassa tarkastellaan betajakauman käyttöä havaintojen jakaumamallina ja erityisesti sen parametrien estimointia maksimointimenetelmällä. Ilmeisesti laskennallisista vaikeuksista johtuen tätä kysymystä on varsin niukasti tarkasteltu alan julkaisuissa. Tavallisesti suositellaan yksinkertaisempaa, mutta epäluotettavampaa momenttimenetelmää¹⁾.

Betajakauman sovellutuksista suurin osa liittyy järjestetyn otoksen tarkasteluihin. Kuten tunnettua samaa jatkuvaa jakaumaa noudattavien, riippumattomien muuttujien kertymäfunktioilla muunnetut järjestetyt havaintoarvot noudattavat betajakaumaa tunnetuin parametrein. Tähän tulokseen perustuen hahmotellaan luvussa 3 päätösvaihtoihin liittyvien kertymäfunktioiden arvostamiskehikko.

Muita kirjallisuudessa käsiteltyjä betajakaumien sovellutuskohteita on esimerkiksi betajakauman käyttö binomijakauman parametrien priorijakaumana. (Katso Baldessari (4), Anscombe (5) ja Rao (6)).

Hahmotellaan lyhyesti tilanteita, joissa betajakauman soveltaminen havaintoaineistoon voi osoittautua hyödylliseksi.

1. Tarkastellaan ei-negatiivisia havaintoja

x_1, x_2, \dots, x_n , joille voidaan asialoogisin perustein arvioida riittävän suuri yläraja M . Esimerkiksi suhdeasteikon muuttujat (ks. Vasama - Vartia (7) s. 49) ovat tavallisesti tätä tyyppiä.

Soveltamalla muunnosta

1) Esimerkiksi Pearson (2) ja Elderton & Johnson (3). Käytännön vaikeuksiin momenttimenetelmää sovellettaessa viitataan esimerkiksi teoksessa Elderton & Johnson (3) sivulla 95. Kirjoittajat ovat kuitenkin tiukasti momenttimenetelmän kan-

$$(2.1.5) \quad y_i = G(x_i) = \frac{x_i}{M}$$

saadaan välin $(0,1)$ havaintoarvoja y_i , joiden jakaumaa voidaan approksimoida betajakaumalla. Mikäli alkuperäisten havaintojen x_i jakauma on likimain normaali-, gamma- tai χ^2 -jakauma on yhteensopivuus odotettavasti hyvä, koska ko. jakaumat ovat betajakauman rajatapauksia.

2. Tarkastellaan edelleen välin $(0, \infty)$ havaintoarvoja x_1, x_2, \dots, x_n . Mikäli näiden jakauma on likimain F-jakauma saadaan muunnoksella

$$(2.1.6) \quad y_i = G(x_i) = \frac{x_i}{\alpha + x_i}, \text{ jossa } y_i \in (0,1)$$

likimain betajakaumaa noudattavia muuttujia, kun $\alpha > 0$ valitaan sopivasti. Tämä johtuu betajakaumien ja F-jakaumien yksikäsitteisestä vastaavuudesta. Tulokseen liittyvä tulojakumasovellutus on Vartia - Vartia (8).

3. Mikäli havaintoaineisto x_1, x_2, \dots, x_n muodostuu jonkin ilmiön tai ominaisuuden suhteellisista esiintymisfrekvensseistä (esim. määrätyn ikäryhmän suhteellinen osuus eri alueilla) on luontevaa kokeilla betajakaumaa havaintojen teoreettisena mallina. Tällöin ei tarvitse ajatella havaintojen x_i syntyneen riippumattomien toistokokeiden realisaatioina, jolloin niiden teoreettisena mallina käytettäisiin tunnettua binomijakauman muunnosta, vaan niiden empiirinen jakauma voi olla jopa U:n muotoinen.

2.2. TIHEYSFUNKTION KÄYTTÄYTYMINEN

Tiheysfunktion

$$(2.2.1) \quad \beta(x|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} \phi(0 < x < 1)$$

yleinen käyttäytyminen riippuu (kerrannoista vakiota

$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} = \frac{1}{B(a,b)}$ vaille) pisteen $x = 0$ ympäristössä parametrissa a ja pisteen $x = 1$ ympäristössä parametrissa b :

$$(2.2.2) \quad \beta(x|a,b) \approx cx^{a-1}, \quad \text{kun } x \gtrsim 0$$

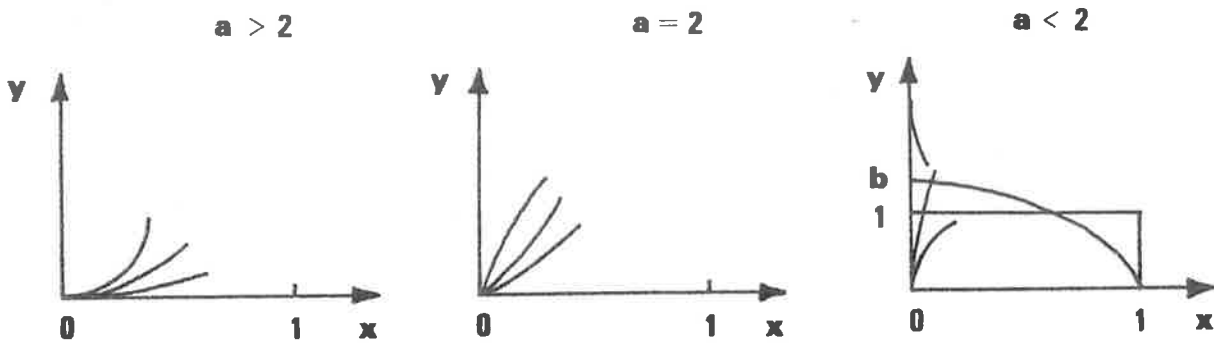
$$(2.2.3) \quad \beta(x|a,b) \approx c(1-x)^{b-1}, \quad \text{kun } x \lesssim 1$$

Symmetrian vuoksi riittää tarkastella tiheysfunktiota origon positiivisessa ympäristössä:

$$(2.2.4) \quad \frac{d}{dx} \beta(x|a,b) \approx \frac{d}{dx} cx^{a-1}$$

$$= c(a-1)x^{a-2}$$

Havaitaan, että tiheysfunktion 'häntä' oskuloi x -akselia origossa, kun $a > 2$, lähestyy origoa positiivisessa kulmassa, kun $a = 2$ ja oskuloi yleensä y -akselia, kun $a < 2$, saaden kuitenkin tapauksessa $a = 1$ origossa arvon b .



Tarkastellaan asiaa tarkemmin, kun $a < 2$.

1. Jos $1 < a < 2$ on $\beta(0|a,b) = 0$ ja lausekkeesta (2.2.4) nähdään, että $\beta(x|a,b)$ oskuloi y-akselia

2. Jos $a = 1$ yksinkertaistuu tiheysfunktio muotoon

$$(2.2.5) \quad \beta(x|1,b) = \frac{\Gamma(1+b)}{\Gamma(1)\Gamma(b)} (1-x)^{b-1}$$

$$= b(1-x)^{b-1},$$

joten origossa tiheysfunktio saa arvon b ja tiheysfunktion derivaatta arvon $-b(b-1)$. Kuviossa $b > 1$.

3. Jos $0 < a < 1$ on $\beta(0|a,b) = \infty$ ja $\beta(x|a,b)$ oskuloi y-akselia. Betajakauman moodi saadaan määrättyä tarkastelemalla tiheysfunktion derivaattaa:

$$(2.2.6) \quad \beta'(x|a,b) = \frac{d}{dx} \left(\frac{x^{a-1}(1-x)^{b-1}}{B(a,b)} \right)$$

$$\propto (a-1)x^{a-2}(1-x)^{b-1} + (b-1)x^{a-1}(1-x)^{b-2}$$

$$= 0, \quad \text{kun } x = 0 \text{ ja } a > 2$$

$$= 0, \quad \text{kun } x = 1 \text{ ja } b > 2$$

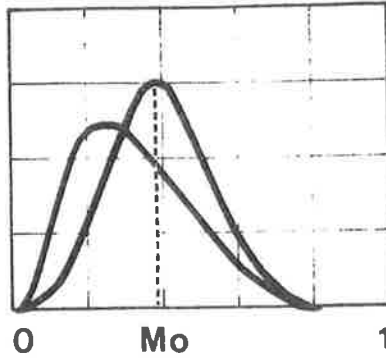
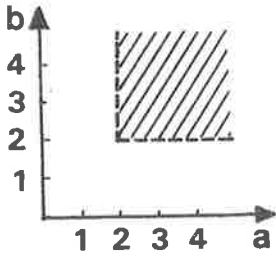
Kolmas mahdollinen nollakohta saadaan, kun $x^{a-2}(1-x)^{b-2} \neq 0$ ja lauseke $(a-1)x^{a-2}(1-x)^{b-1} + (b-1)x^{a-1}(1-x)^{b-2} = 0$, josta ratkaisuna x saadaan moodin arvoksi

$$(2.2.7) \quad Mo = \frac{a-1}{a+b-2}$$

Tällöin on oletettava, että a ja $b > 1$. Jos a tai $b \leq 1$ on moodina 0 tai 1. Tapauksessa $a = b = 1$ (tasajakauma) ei moodia ole.

Seuraavassa on esitetty tyypillisiä tapauksia graafisesti. Vasemman laidan kuvio osoittaa parametrien a ja b vaihtelualueen; mikäli reuna on piirretty yhtenäisellä viivalla kuuluvat reunapisteet mukaan. Keskellä olevassa kuviossa esitetään tyypillisiä tiheysfunktioita. Parametrin θ määrittely on suoritettu sivulla 19.

$a \text{ ja } b > 2$

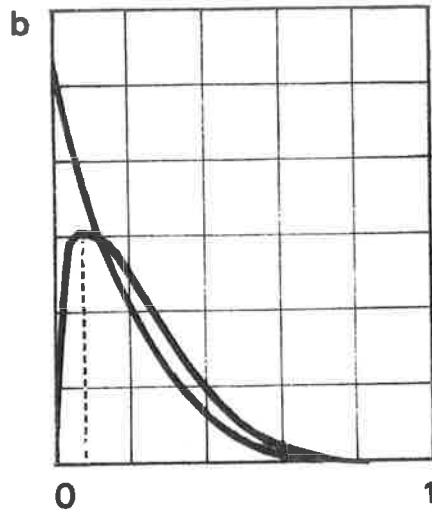
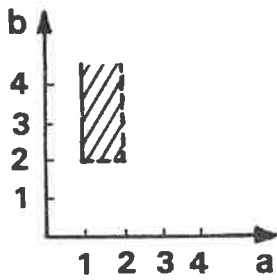


$$Mo = \frac{a-1}{a+b-2}$$

$$\theta = \frac{1}{a+b+1} < \frac{1}{5}$$

$1 \leq a < 2$

$b > 2$

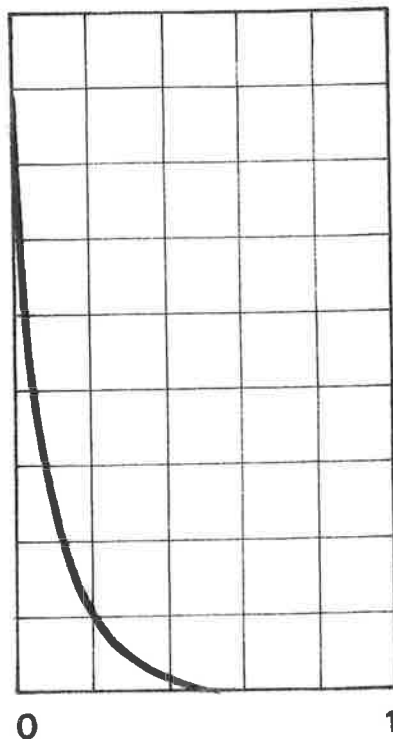
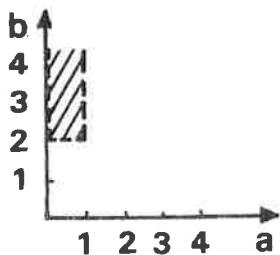


$$Mo = \frac{a-1}{a+b-2}$$

$$\theta = \frac{1}{a+b+1} < \frac{1}{4}$$

$0 < a < 1$

$b > 2$

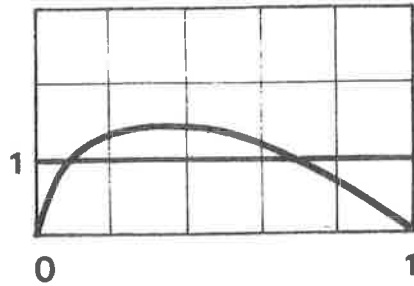
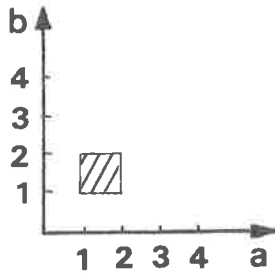


$$Mo = 0$$

$$\theta < \frac{1}{3}$$

$1 \leq a \leq 2$

$1 \leq b \leq 2$

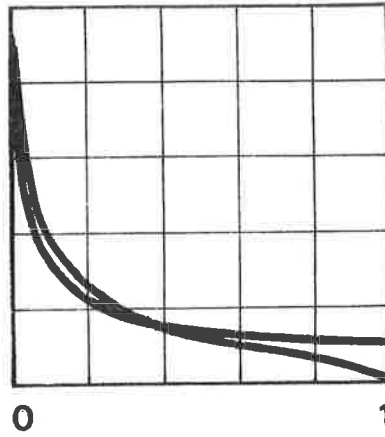
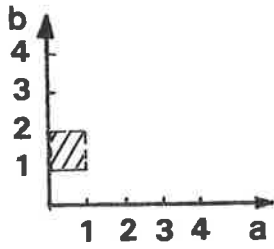


$M_0 = \frac{a-1}{a+b-2}$

$\frac{1}{5} < \theta < \frac{1}{3}$

$0 < a < 1$

$1 \leq b < 2$

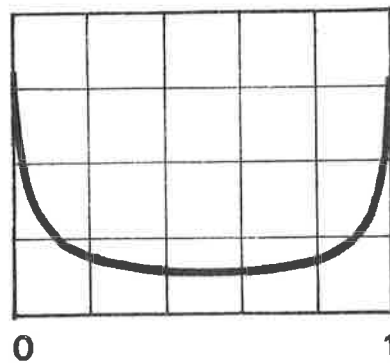
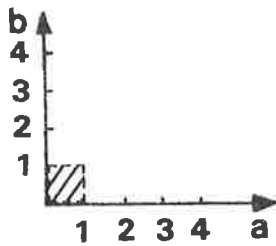


$M_0 = 0$

$\frac{1}{4} < \theta < \frac{1}{2}$

$0 < a < 1$

$0 < b < 1$



$M_0 = 0$ tai

$M_0 = 1$

$\frac{1}{3} < \theta < 1$

2.3. TUNNUSLUVUT

Seuraavassa esitellään luettelonomaisesti betajakauman tavallisimmat tunnusluvut ja eräitä niiden ominaisuuksia; johtoja ei suoriteta, koska ne löytyvät alan oppikirjoista.

(2.3.1) n:s origomomentti

$$\alpha_n = E\underline{x}^n = \frac{\Gamma(a+n)\Gamma(a+b)}{\Gamma(a)\Gamma(a+b+n)}$$

Käyttämällä merkintöjä $x = a+b$ ja $(x)_n = \frac{\Gamma(x+n)}{\Gamma(x)}$

(ns. Pochhammerin symboli) saadaan

$$(2.3.2) \quad \alpha_n = \frac{(a)_n}{(x)_n} = \frac{a(a+1)\dots(a+n-1)}{x(x+1)\dots(x+n-1)}$$

jossa jälkimmäinen yhtälö pätee vain kun $n \in \mathbb{N}$.

Erityisesti on

(2.3.3) odotusarvo

$$\alpha_1 = E\underline{x} = \frac{a}{a+b} = \text{df } p$$

(2.3.4) 2. origomomentti

$$\alpha_2 = E\underline{x}^2 = \frac{a(a+1)}{x(x+1)} = \alpha_1 \left(\frac{a+1}{x+1}\right)$$

ja yleisesti $\alpha_{n+1} = \alpha_n \left(\frac{a+n}{x+n}\right)$.

(2.3.5) Varianssi

$$D^2 \underline{x} = \frac{ab}{(a+b)^2 (a+b+1)} = \frac{pq}{x+1}, \text{ jossa } q = 1-p$$

Varianssi $D^2 \underline{x}$ on kätevä esittää muodossa

$$(2.3.6) \quad D^2 \underline{x} = pq \theta, \text{ jossa } \theta = \frac{1}{a+b+1} = \frac{1}{x+1}$$

Parametri θ voidaan tulkita betajakauman ja binomijakauman

$\text{Bin}(1,p)$ välisen 'sukulaisuuden' mitaksi: binomijakauma

$\text{Bin}(1,p)$ on singulaarinen betajakauma parametrien arvoilla

$\theta = 1$ ja $p = \lim_{a,b \rightarrow 0} \left(\frac{a}{a+b}\right)$. Parametrien p ja θ käyttö on

miellyttävää, koska molemmat kuuluvat väliin $[0,1]$. Vastaavasti

a ja b voidaan lausua p :ssä ja θ :ssa seuraavasti:

$$(2.3.7) \quad a = p \left(\frac{1-\theta}{\theta}\right)$$

$$(2.3.8) \quad b = (1-p) \left(\frac{1-\theta}{\theta}\right) = q \left(\frac{1-\theta}{\theta}\right)$$

(2.3.9) Variaatiokerroin

$$v = \frac{Dx}{Ex} = \sqrt{\frac{b}{a} \left(\frac{1}{a+b+1}\right)}$$

(2.3.10) Pearsonin vinouskerroin

$$\sqrt{\beta_1} = \frac{\mu_3}{\sigma^3} = \frac{2(b-a) \sqrt{a+b+1}}{(a+b+2) \sqrt{ab}}$$

(2.3.11) Pearsonin huipukkuuskerroin

$$\beta_2 = \frac{\mu_4}{\sigma^4} = \frac{3(a+b+1)[2(a+b)^2 + ab(a+b-6)]}{ab(a+b+2)(a+b+3)}$$

(2.3.12) Moodi

$$Mo = \frac{a-1}{a+b-2}, \quad a > 1$$

(2.3.13) Harmoninen keskiarvo

$$H = \frac{a-1}{a+b-1}$$

Tämä saadaan lausekkeesta $\frac{1}{H} = E\left(\frac{1}{X}\right)$.

(2.3.14) Geometrinen keskiarvo

$$f = \text{Geom}(\underline{x}) = \frac{e^{\psi(a)}}{e^{\psi(x)}}$$

jossa $\psi(x) = \frac{d}{dx} \log \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}$ (ns. psii-funktio).

Tämä saadaan lausekkeesta

$$(2.3.15) \quad \log f = E \log \underline{x} = \int_0^1 \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1} \log x \, dx$$

käyttämällä tulosta $\frac{\partial}{\partial a} B(a,b) = \int_0^1 x^{a-1} (1-x)^{b-1} \log x \, dx$.

Tunnuslukua f sanotaan myös alhaalta lasketuksi geometriseksi keskiarvoksi erotukseksi ylhäältä lasketulle geometriselle keskiarvolle $g = \text{Geom}(1-\underline{x})$, joka on myös jakauman $\text{Beta}(b, a)$ (alhaalta laskettu) geometrinen keskiarvo.

Seuraavassa tullaan käyttämään myös suureita

$$(2.3.16) \quad u =_{df} E \log\left(\frac{1}{\underline{x}}\right) = \log\left(\frac{1}{\underline{f}}\right) = \psi(a+b) - \psi(a)$$

$$(2.3.17) \quad v =_{df} E \log\left(\frac{1}{1-\underline{x}}\right) = \log\left(\frac{1}{\underline{g}}\right) = \psi(a+b) - \psi(b),$$

joista esim. u ilmaisee montako senttidyniä¹⁾ suurempi 1 on geometrinen keskiarvo f .

Lause:
$$v = \log\left(\frac{1}{\underline{g}}\right) = \psi(a+b) - \psi(b) = \sum_{k=1}^{\infty} \frac{\alpha_k}{k}$$

tod.
$$\psi(a+b) - \psi(b) = - E \log(1-\underline{x})$$

$$= E\left(\sum_{k=1}^{\infty} \frac{\underline{x}^k}{k}\right)$$

$$= \sum E\left(\frac{\underline{x}^k}{k}\right)$$

$$= \sum \left(\frac{\alpha_k}{k}\right) \quad \square$$

1) Suhteen $\frac{x}{y}$ 100-kertainen luonnollinen logaritmi, $100 \log\left(\frac{x}{y}\right)$, ilmaisee montako dyniä suurempi x on y :tä. Pienillä erojen arvoilla tämä dynimuutos vastaa prosenttimuutosta, koska $\log \frac{x}{y} = \log\left(1 + \frac{x-y}{y}\right) \approx \frac{x-y}{y}$.

Vastaavasti $\log\left(\frac{1}{f}\right) = \sum_{k=1}^{\infty} \left(\frac{\beta_k}{k}\right)$, jossa $\beta_k = E(1-x)^k$ on

Beta(b,a):n k:s origomomentti. Asettamalla $a=1$ psii-funktion lisäyslausekkeessa

$$(2.3.18) \quad \psi(a+b) - \psi(b) = \sum_{k=1}^{\infty} \frac{\alpha_k}{k}$$

saadaan klassinen tulos (katso esim. Abramowitz-Stegun () sivu 259)

$$(2.3.19) \quad \psi(b+1) - \psi(1) = \sum_{k=1}^{\infty} \frac{b}{k(b+k)}$$

Lause: $0 \leq f + g \leq 1$

tod. koska $0 \leq f \leq E\underline{x}$ ja $0 \leq g \leq E(1-\underline{x})$

on $0 \leq f + g \leq E\underline{x} + E(1-\underline{x}) = 1$. □

Koska geometrinen keskiarvo on aina aritmeettisen ja harmoni-
sen keskiarvon välissä, $H \leq f \leq E\underline{x}$, saadaan f:lle epäyhtälö

$$(2.3.20) \quad \frac{a-1}{a+b-1} \leq f \leq \frac{a}{a+b}$$

Alaraja on mielekäs vain kun $a > 1$, tarkkaa f:n arviota ei tästä saada kuin suurilla a:n ja b:n arvoilla.

2.4. MAXIMUM LIKELIHOOD-YHTÄLÖT: KIRJOITUSTAPOJA JA
TULKINTOJA

Siirrytään tarkastelemaan n -ulotteista riippumatonta otosta $\underline{X} = (\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n)$, jossa $\underline{x}_i \sim \text{Beta}(a, b)$. Likelihoodfunktion logaritmi $l(X|a, b)$ saa kiinteässä otoksessa $X = (x_1, x_2, \dots, x_n)$ esityksen

$$\begin{aligned}
 (2.4.1) \quad l(X|a, b) &= \sum_{i=1}^n \log f(x_i|a, b) \\
 &= \sum_{i=1}^n \log \beta(x_i|a, b) = \sum_{i=1}^n \log \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x_i^{a-1} (1-x_i)^{b-1} \\
 &= n\{\log \Gamma(a+b) - \log \Gamma(a) - \log \Gamma(b)\} \\
 &\quad + (a-1) \sum_{i=1}^n \log x_i + (b-1) \sum_{i=1}^n \log (1-x_i)
 \end{aligned}$$

Likelihood yhtälöiksi saadaan siis

$$(2.4.2) \quad \frac{\partial l}{\partial a} = n\left\{\frac{\partial}{\partial a} \log \Gamma(a+b) - \frac{\partial}{\partial a} \log \Gamma(a)\right\} + \sum_{i=1}^n \log x_i = 0$$

$$(2.4.3) \quad \frac{\partial l}{\partial b} = n\left\{\frac{\partial}{\partial b} \log \Gamma(a+b) - \frac{\partial}{\partial b} \log \Gamma(b)\right\} + \sum_{i=1}^n \log (1-x_i) = 0$$

Käyttämällä hyväksi alueella $0 < x < \infty$ aidosti kasvavaa psii-funktiota¹⁾

$$(2.4.4) \quad \psi(x) = \frac{d}{dx} \log \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}$$

1) Psii-funktiolle käytetään myös nimitystä 'digammafunktio' ja joskus symbolia $\Psi(x)$. Valitettavasti sekaannusta syntyy, kun erityisesti saksalaisessa kirjallisuudessa sovelletaan samaa symbolia Ψ funktiolle $\psi(x+1) = \psi(x) + \frac{1}{x}$. Horjuvuus merkinöissä palautuu vanhaan kiistaan Γ -funktion määrittelystä (vrt. esim. Fulks (10) s. 452).

ja seuraavia merkintöjä otoksesta lasketuille suureille

$$(2.4.5) \quad \hat{u} = \frac{1}{n} \sum_{i=1}^n \log\left(\frac{1}{x_i}\right) \geq 0$$

$$(2.4.6) \quad \hat{v} = \frac{1}{n} \sum_{i=1}^n \log\left(\frac{1}{1-x_i}\right) \geq 0, \quad ,$$

voidaan likelihood yhtälöt lausua seuraavassa symmetrisessä muodossa

$$(2.4.7) \quad \psi(a+b) - \psi(a) = \hat{u}$$

$$(2.4.8) \quad \psi(a+b) - \psi(b) = \hat{v}.$$

Määritellään ns. alhaalta laskettu geometrinen keskiarvo \hat{f} ja ylhäältä laskettu geometrinen keskiarvo \hat{g} yhtälöillä

$$(2.4.9) \quad \hat{f} = \left(\prod_{i=1}^n x_i \right)^{\frac{1}{n}}$$

$$(2.4.10) \quad \hat{g} = \left(\prod_{i=1}^n (1-x_i) \right)^{\frac{1}{n}}, \quad 0 < x_i < 1.$$

Otossuureiden \hat{u} ja \hat{v} sekä \hat{f} ja \hat{g} välillä on seuraavat täsmäl-
liset yhteydet

$$(2.4.11) \quad \hat{u} = \log\left(\frac{1}{\hat{f}}\right), \quad , \quad \hat{f} = e^{-\hat{u}}$$

$$(2.4.12) \quad \hat{v} = \log\left(\frac{1}{\hat{g}}\right), \quad , \quad \hat{g} = e^{-\hat{v}}.$$

On siis osoitettu, että betajakauman parametrien a ja b maximum likelihood-estimaattorit \hat{a} ja \hat{b} ovat otoksen (x_1, x_2, \dots, x_n) alhaalta ja ylhäältä laskettujen geometrinen keskiarvojen \hat{f} ja \hat{g} funktiota.

Maximum likelihood-menetelmä asettaa siis betajakauman tapauksessa seuraavan vastaavuuden teoreettisten ja empiiristen suureiden välille

$$(2.4.13) \quad \text{est } u = \hat{u} = \frac{1}{n} \sum \log\left(\frac{1}{x_i}\right)$$

$$(2.4.14) \quad \text{est } v = \hat{v} = \frac{1}{n} \sum \log\left(\frac{1}{1-x_i}\right)$$

Maximum likelihood-menetelmä voidaankin tulkita analogiamenetelmäksi, jossa teoreettiset geometrinen keskiarvot $f = \text{Geom}(\underline{x}) = e^{-u}$ ja $g = \text{Geom}(1-\underline{x}) = e^{-v}$ asetetaan yhtä suuriksi vastaavien otoskeskiarvojen $\hat{f} = (\prod x_i)^{\frac{1}{n}}$ ja $\hat{g} = (\prod (1-x_i))^{\frac{1}{n}}$ kanssa ja ratkaistaan parametrit a ja b näin syntyvästä yhtälöryhmästä:

$$(2.4.15) \quad \psi(a+b) - \psi(a) = u = \hat{u}$$

$$(2.4.16) \quad \psi(a+b) - \psi(b) = v = \hat{v}$$

Tällä yhtälöryhmällä on yksikäsitteinen ratkaisu, kun \hat{u} ja \hat{v} ovat positiivisia (eli jokainen $x_i > 0$) ja

$$(2.4.17) \quad \hat{f} + \hat{g} = e^{-\hat{u}} + e^{-\hat{v}} \leq 1 .$$

Tämä epäyhtälö pätee aina teoreettisille geometrisille keskiarvoille f ja g , kuten on sivulla 22 osoitettu. Mutta se pätee myös empiirisille geometrisille keskiarvoille, sillä

$$(2.4.18) \quad \hat{f} \leq \bar{x} \quad , \quad \hat{g} \leq 1 - \bar{x} \quad , \quad \text{joten}$$

$$(2.4.19) \quad \hat{f} + \hat{g} \leq \bar{x} + (1 - \bar{x}) = 1 \quad .$$

Keskiarvon \hat{g} epäyhtälö seuraa siitä, että

$$(2.4.20) \quad \hat{g} = (\prod (1-x_i))^{1/n} \leq M(1-x) = \frac{1}{n} \sum (1-x_i) = 1 - \bar{x} \quad .$$

Tämä osoittaa, että havaittujen empiiristen geometrinen keskiarvojen \hat{f} ja \hat{g} suuruiset teoreettiset keskiarvot f ja g ovat aina olemassa, joten ne voidaan merkitä yhtä suuriksi. Näin ollen ainoana ongelmana on yhtälöryhmän (2.4.15-16) ratkaiseminen.

Seuraavassa tulkitsemme nämä yhtälöt ainoastaan teoreettisten suureiden väliseksi yhteyksiksi, jolloin voimme puhua vain parametreista tai teoreettisista tunnusluvuista otossuureiden ja estimaattoreiden asemasta. Voimme tavallaan jättää sivuun varsinaisen estimointiongelman ja siirtyä tarkastelemaan betajakauman tunnuslukujen f ja g (tai u ja v) ja parametrien a ja b välistä kääntäen yksikäsitteistä riippuvuutta. Tämä käytäntö selventää esitystä ja antaa sille yleisempääkin kiinnostavuutta.

Ratkaistavat yhtälöt, joita kuitenkin vielä kutsumme maximum likelihoid yhtälöksi, voidaan esittää eri tavoilla käyttökelpoisissa muodoissa

$$(2.4.21) \quad \underline{\text{Perusmuoto}} \quad \begin{aligned} \psi(a+b) - \psi(a) &= u \\ \psi(a+b) - \psi(b) &= v \end{aligned}$$

Käyttämällä merkintöjä

$$(2.4.22) \quad x = a+b \quad , \quad 0 < x < \infty$$

$$(2.4.23) \quad y = \psi(x) = \psi(a+b) \quad , \quad -\infty < y < \infty$$

$$(2.4.24) \quad x = \psi^{-1}(y) = \text{psiin käänteisfunktio}$$

saadaan seuraava ns. käänteisfunktio muoto.

$$(2.4.25) \quad \underline{\text{Käänteisfunktio muoto}}$$

$$\psi^{-1}(y) = \psi^{-1}(y-u) + \psi^{-1}(y-v)$$

Tämä esityksen etuna on ongelman palantuminen yhden tuntemattoman $y = \psi(x)$ (tai sen sopivan funktion kuten $z = e^y$) yhtälöön. Kun yhtälön (2.4.25) ratkaisu y tunnetaan saadaan a ja b yhtälöistä

$$(2.4.26) \quad a = \psi^{-1}(y-u)$$

$$(2.4.27) \quad b = \psi^{-1}(y-v)$$

jotka ovat juuri (2.4.25):n termejä. Käänteisfunktio-
muodon soveltamista varten on liitteessä 1 johdettu psiin käänteis-
funktion approksimaatiokaavoja.

Käänteisfunktio-
muotoa sovelletaan, kun a ja b molemmat ovat
suuria. Kun a ja b ovat pieniä, saadaan likelihood yhtälöille
seuraavanlaiset käyttökelpoiset esitykset. Käyttämällä psi-
funktion siirto-ominaisuutta

$$(2.4.28) \quad \psi(x+1) = \psi(x) + \frac{1}{x} ,$$

joka on seuraus gamma-funktion kertomaominaisuudesta
 $\Gamma(x+1) = x\Gamma(x)$, ja määrittelemällä suureet

$$(2.4.29) \quad u' = \psi(a+b+1) - \psi(a+1) \geq 0$$

$$(2.4.30) \quad v' = \psi(a+b+1) - \psi(b+1) \geq 0 ,$$

saadaan likelihood yhtälöiden ns. siirtomuoto.

$$(2.4.31) \quad \underline{\text{Siirtomuoto}}$$

$$\frac{b}{a(a+b)} = u - u' = df \bar{u}$$

$$\frac{a}{b(a+b)} = v - v' = df \bar{v}$$

Tunnusluku u' on jakauman $\text{Beta}(a+1, b)$ alhaalta lasketun

geometrisen keskiarvon logaritmi, $u' = -\log f'$, ja v' on vastaavasti $\text{Beta}(a, b+1)$:n ylhäältä lasketun geometrisen keskiarvon logaritmi, $v' = \log g'$. Tunnusluku $\bar{u} = u - u' = \log(f/f')$ ilmoittaa montako senttidyniä alhaalta laskettu geometrinen keskiarvo muuttuu kun b pysyy vakiona ja a kasvaa yhdellä; samoin $\bar{v} = v - v' = \log(g/g')$ kertoo saman senttidynimuutoksen ylhäältä lasketulle geometriselle keskiarvolle, kun a on vakio ja b kasvaa yhdellä.

Kertomalla ja jakamalla siirtomuodon yhtälöt keskenään ja ottamalla neliöjuuret saadaan suureiden $x = a+b$ ja $p/q = a/b$ määrittämisenä käyttökelpoiset yhtälöt

$$(2.4.32) \quad \frac{1}{x} = \sqrt{(u - u')(v - v')} = \sqrt{uv}$$

$$(2.4.33) \quad \frac{a}{b} = \sqrt{\frac{v - v'}{u - u'}} = \sqrt{\frac{v}{u}} .$$

Koska $u' \rightarrow 0$ ja $v' \rightarrow 0$, kun a ja $b \rightarrow 0$, saadaan seuraavat arviot, jotka myös perustelevat käytettyjä merkintöjä

$$(2.4.34) \quad \frac{1}{x} = \frac{1}{a+b} \approx \sqrt{uv}$$

$$(2.4.35) \quad \frac{a}{b} = \frac{p}{q} \approx \sqrt{\frac{v}{u}} .$$

Perusmuoto voidaan kirjoittaa myös muotoon

$$(2.4.36) \quad \psi\left(\frac{1-\theta}{\theta}\right) - \psi\left(p\left(\frac{1-\theta}{\theta}\right)\right) = u$$

$$(2.4.37) \quad \psi\left(\frac{1-\theta}{\theta}\right) - \psi\left(q\left(\frac{1-\theta}{\theta}\right)\right) = v$$

käyttämällä suureita $\theta = \frac{1}{a+b+1}$, $p = \frac{a}{a+b}$ ja $q = 1 - p$.

Tämän esityksen avulla on laskettu liitetaulukko 4, joka sisältää parametrin θ arvoja 0.1, 0.2, ..., 0.9 ja parametrin p arvoja 0.1, 0.2, ..., 0.5 vastaavat parametrien a , b , u , v , f , g arvot. Taulukon avulla on kätevä tutkia kehitettävien approksimaatioiden tarkkuuksia.

2.5. BETAJAKAUMAN JA ERÄIDEN MUIDEN JAKAUMIEN VÄLISIÄ YHTEYKSIÄ

Jos betajakaumassa $Beta(a,b)$ pidetään odotusarvo $p = \frac{a}{a+b}$ vakiona, kun a ja b kasvavat, lähestyy jakaumaa $Beta(a,b)$ noudattava satunnaismuuttuja \underline{x} vakiota p . Tämä lähestyminen tapahtuu siten, että normeeratun muuttujan $\underline{z} = \frac{\underline{x} - p}{\sigma}$ jakauma lähestyy normaalijakaumaa $N(0,1)$. Väitteen todistus on löydettävissä useista alan oppikirjoista. Edellä on jo viitattu siihen, että $Beta(a,b)$ lähenee binomijakaumaa $Bin(n=1, p = \frac{a}{a+b})$, kun a ja b lähestyvät nollaa.

Jos parametria a pidetään kiinteänä ja b kasvaa rajatta, lähestyy muuttujan $\underline{y} = b\underline{x}$ jakauma gammajakaumaa parametrein $(1,a)$. Esimerkiksi Cramér (11) soveltaa tätä seikkaa järjestettyyn otokseen liittyvien satunnaismuuttujien tutkimisessa. Tämän perusteella voidaan hyvin vinojen betajakaumien parametrien estimoinnissa käyttää gammajakauman ominaisuuksia tai käyttää gammajakaumaa sopivasti muunnettujen havaintojen teoreettisena jakaumamallina.

Esitellään betajakauman yhteydet χ^2 , gamma- ja F-jakaumiin. Seuraavassa luetellaan ko. jakaumien tiheysfunktiot; merkintä $\emptyset(A)$ tarkoittaa joukon A indikaattorifunktiota.

Betajakauma parametrein (a,b) :

$$(2.5.1) \quad \beta(x|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} \emptyset(0 < x < 1).$$

χ^2 -jakauma vapausasteella k:

$$(2.5.2) \quad \chi^2(x|k) = \frac{1}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})} e^{-\frac{x}{2}} x^{\frac{k}{2}-1} \varnothing(0 < x < \infty).$$

Tämä on erikoistapaus gammajakaumasta parametrein (α, λ) :

$$(2.5.3) \quad \gamma(x|\alpha, \lambda) = \frac{\alpha^\lambda}{\Gamma(\lambda)} e^{-\alpha x} x^{\lambda-1} \varnothing(0 < x < \infty).$$

F-jakauma vapausastein (m, n) :

$$(2.5.4) \quad F(x|m, n) = \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})} \left(\frac{m}{n}\right)^{\frac{m}{2}} x^{\frac{m}{2}-1} \left(1+\frac{m}{n}x\right)^{-\frac{m+n}{2}} \varnothing(0 < x < \infty).$$

Esimerkiksi Rao (6) ja Cramér (11) esittävät seuraavat yhteydet näiden jakaumien välillä:

1. Jos \underline{x} ja \underline{y} ovat riippumattomia gammamuuttujia parametrein (α, λ_1) ja (α, λ_2) on

$$(2.5.5) \quad \underline{g} = \frac{\underline{x}}{\underline{x}+\underline{y}}$$

betamuuttuja parametrein (λ_1, λ_2) .

2. Olkoon \underline{x} F-muuttuja vapausastein (m, n) . Tällöin on

$$(2.5.6) \quad \underline{g} = \frac{\underline{x}}{\left(\frac{n}{m}\right)+\underline{x}}$$

betamuuttuja parametrein $\left(\frac{m}{2}, \frac{n}{2}\right)$.

Vastaava tulos on $g:n$ ja $x:n$ välisen yksikäsitteiden yhteyden vuoksi voimassa myös toiseen suuntaan. Olkoon g beta-muuttuja parametrein $(\frac{m}{2}, \frac{n}{2})$. Tällöin on

$$(2.5.7) \quad \underline{x} = \binom{n}{m} \left(\frac{g}{1-g} \right)$$

F-muuttuja vapausastein (m,n) . Tämän nojalla betajakaumat ja F-jakaumat vastaavat yksikäsitteisesti toisiaan, mikäli F-jakauman vapausasteiden ajatellaan saavan kaikki positiiviset reaaliarvot.

3. Olkoon jatkuvan satunnaismuuttujan u tiheysfunktio $f(u)$ ja kertymäfunktio $F(u)$. Jos u :lle suoritetaan seuraava todennäköisyysintegraalimuunnos

$$(2.5.8) \quad \underline{x} = \int_{-\infty}^{\underline{u}} f(t) dt = F(\underline{u}),$$

eli jos u :ta muunnetaan sen omalla kertymäfunktioilla, saadaan generoitua tasaisesti välillä $(0,1)$ jakautunut satunnaismuuttuja \underline{x} eli $\underline{x} \sim \text{Beta}(1,1)$.

Olkoon nyt $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ riippumattomia tasaisesti välillä $(0,1)$ jakautuneita muuttujia. Tällöin noudattavat muuttujat

$$(2.5.9) \quad \underline{y}_i = -2 \log \underline{x}_i$$

gammajakaumaa parametrein $(\frac{1}{2}, 1)$ eli χ^2 -jakaumaa vapausasteella

2 ja ovat myös riippumattomia. χ^2 -jakauman yhteenlaskuominaisuuden nojalla

$$(2.5.10) \quad \underline{Y} = Y_1 + Y_2 + \dots + Y_n = -2\log \underline{x}_1 - 2\log \underline{x}_2 - \dots - 2\log \underline{x}_n$$

on χ^2 -muuttuja vapausasteella $2n$. Tämän tuloksen perusteella voidaan kombinoida useita riippumattomia testejä, kuten esim. Rao (6) sivulla 137 mainitsee.

3. ERÄS SOVELLUTUSALUE: PÄÄTÖSVAIHTOEHTOJEN
PAREMMUUSJÄRJESTYKSEN MÄÄRÄÄMINEN

Olkoon $F(x)$ jatkuva kertymäfunktio eli ns. kantafunktio.
Tarkastellaan kertymäfunktioiperhettä, jonka elementit voidaan
esittää muodossa

$$(3.1) \quad G(x) = \int_{-\infty}^x F(x)^{a-1} (1-F(x))^{b-1} dF(x) \\ = \int_0^{F(x)} y^{a-1} (1-y)^{b-1} dx$$

Tämä merkitsee, että tarkastellaan satunnaismuuttujaa \underline{x} , joka
kantafunktiolla $F(x)$ muunnettuna noudattaa betajakaumaa:

$$(3.2) \quad F(\underline{x}) \sim \text{Beta}(a, b)$$

Niiden satunnaismuuttujien \underline{x} joukko, joiden kertymäfunktio
 $G(x)$ voidaan esittää muodossa (3.1) tai vaihtoehtoisesti, joi-
den muunnoksille $F(\underline{x})$ pätee (3.2), on erittäin laaja ja sisäl-
tää hyvin eri tavalla jakautuneita muuttujia. Näin ollen mene-
telmiä, joita kehitetään tälle muuttujajoukolle, voidaan so-
veltaa mitä erilaisimpiin ongelmiin. Käytännössä ei ole tar-
peen, että oletukset (3.1) tai (3.2) ovat tarkasti voimassa,
vaan kehitettävien menetelmien soveltamiseksi riittää oletuk-
sien likimääräinen toteutuminen.

Esimerkkinä tarkasteltavan muuttujajoukon laajuudesta esite-
tään seuraava tärkeä tapaus, jossa oletukset ovat täsmälleen

voimassa. Olkoot muuttujat

$$(3.3) \quad Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}$$

jakaumasta $F(x)$ poimitun $n:n$ suuruisen otoksen havainnot suuruusjärjestyksessä (pienimmästä suurimpaan). Mikäli alkuperäisen perusjoukon kertymäfunktio $F(x)$ on jatkuva, on jokaisen järjestetyn otoksen (3.3) havainnon kertymäfunktio tyyppiä (3.1) kaikilla otoskoon arvoilla. (Hyvä esitys on Wilks (12) s. 234).

Tarkemmin ilmaistuna on $n:n$ suuruisen otoksen $k:n$ nen havainnon $\underline{x} = Y_{(k)}$ jakauma tyyppiä (3.1) parametrien ollessa $a = k$ ja $b = n-k+1$. Tämä on yhtäpitävää sen kanssa, että kertymäfunktioilla $F(x)$ välille $(0,1)$ muunnetun havainnon $\underline{x} = Y_{(k)}$ jakauma on betajakauma parametrein $a = k$ ja $b = n-k+1$ eli

$$(3.4) \quad F(\underline{x}) \sim \text{Beta}(k, n-k+1).$$

Tarkasteltava muuttujien \underline{x} joukko muodostuu tässä tapauksessa kaikkien eri otoskokojen n otoksista määräytyistä järjestetyistä havainnoista.

Voimmekin pitää tulosta (3.4) betajakauman oleellisena sisältönä: $\text{Beta}(a, b)$ on tietyn kokoisesta otoksesta määrätyn todennäköisyysasteikolle muunnetun järjestetyn havainnon jakauma, jos a ja b ovat kokonaislukuja. Kokonaisluvuista poikkeaviin parametriarvoihin liittyvät jakaumat saadaan ns. analyttisen jatkuvuuden periaatteella.

On hyödyllistä tulkita betajakaumaan liittyvät käsitteet ja operaatiot tuloksen (3.4) avulla, jolloin jakaumasta saadaan ilmeikkäämpi ja jos mahdollista konkreettinen käsitys.

Esimerkiksi parametrien summa $x = a+b$ vastaa otoskokoa lisätynä yhdellä eli $x = k+(n-k+1) = n+1$. Parametrit a ja b vastaavat järjestetyn otoksen havainnon järjestyslukuja lasketuna toisaalta alhaalta ja ylhäältä eli $a = k$ ja $b = n-k+1$. On mielenkiintoista, että nämä tulkinnot voidaan suorittaa vain kun k ja n ovat molemmat vähintään yksi, jolloin myös a ja b ovat vähintään yksi. Tämä sulkee pois kaikki $U:n$ muotoiset betajakaumat, joihin liittyvät tarkastelut ovatkin aivan omalaatuisia.

Kehitettävät menetelmät ovat siinä mielessä jakaumista vapaita, että kantafunktio $F(x)$ saa olla mikä tahansa kiinteä, mutta mielivaltainen jatkuva kertymäfunktio. Mikäli tarkastellaan vain välille $(0,1)$ muunnettuja havaintoja, joka hyvin monissa ongelmissa on mahdollista, tarvitaan kertymäfunktioita $F(x)$ vain muunnoksessa $\underline{x} \rightarrow F(\underline{x})$, jonka jälkeen tarkastelu on siitä riippumaton.

Törnqvistin ja Nordbergin (13) tarkastelema arvostustulosten jakaumaperhe, ns. kantafunktion omaava arvostuskenttä, muodostui tyyppiä

$$(3.5) \quad G(x) = F(x)^a, \quad a > 0$$

olevista kertymäfunktioista, joissa kantafunktio $F(x)$ on jatkuva kertymäfunktio. Tarkastellulla kentällä on useita mielenkiintoisia ominaisuuksia. Esimerkiksi se on suhdestabiili.

Osoitamme, että (3.5) on erikoistapaus (3.1):stä.

Lause 1: Jos \underline{x} :n kertymäfunktio $G(x)$ on yhtälön (3.5) ilmaisemaa tyyppiä, niin $F(\underline{x}) \sim \text{Beta}(a,1)$.

Tod.
$$P(\underline{x} \leq x) = F(x)^a = \int_0^{F(x)} \left(\frac{1}{a-1}\right) y^{a-1} dy$$

$$\Leftrightarrow F(\underline{x}) \sim \text{Beta}(a,1) \quad \square$$

Lauseen perusteella \underline{x} :n jakauma (3.5) voidaan tulkita a :n suuruisesta otoksesta määrätyn suurimman (eli a :nnen) havainnon $\underline{x}_{(a)}$ jakaumaksi, jos $a \in \mathbb{N}$.

Tässä on syytä olla tarkkana järjestyksen suunnasta. Jos suuria arvostustuloksia pidetään hyvinä on suurin arvostustulos paras. Tätä havaintoa voitaisiin kutsua voittajaksi tai (hieman epämääräisesti) "ensimmäiseksi". Tavallisesti kuitenkin järjestettyjen otosten tarkasteluissa "ensimmäiseksi" sanotaan pienintä havaintoa, jolloin suurin havainto n :n suuruudessa otoksessa on n :s havainto. Tämä n :s havainto on siis viimeinen havainto suuruusjärjestyksessä pienimmästä suurimpaan. Tällöin havainnon järjestysluku ja arvostustulos kasvavat samaan suuntaan. Kääntämällä järjestys suurimmasta pienimpään kasvavat järjestysluvut ja arvostustulokset eri suuntiin.

Kumpaakin käytäntöä sovelletaan esim. urheilukilpailuissa. Tämä on ehkä kehittynein instituutio, jossa järjestyksen määrittäminen on perusongelmana. Toinen hyvä esimerkki on äänestystilanne, jossa usein syntyy sekavuutta eri käytäntöjen vuoksi: ovatko ehdokkaille annettavat pisteet järjestyslukuja vai arvostuspisteitä (ääniä)?

Törnqvistin ja Nordbergin (13) tarkastelemassa tilanteessa voidaan siis havaintojen jakaumat tulkita $n:n$ suuruisen otoksen suurimman havainnon jakaumaksi. On ilmeistä, että näiden jakaumien välillä on yksikäsitteinen ja luonnollinen preferenssijärjestys: aina suuremmasta otoksesta saatua suurinta havaintoa pidetään parempana kuin pienemmästä otoksesta saatua. Varsinaisia punnitus- tai arvostusongelmia ei siis tällä tasolla (3.5):n tapauksessa esiinny.

Yleisempi betajakaumaan liittyvä arvostuskenttä, jossa

$$(3.6) \quad F(\underline{x}) \sim \text{Beta}(a, b)$$

on tässä suhteessa huomattavasti rikkaampi ja mielenkiintoisempi.

Betajakauma voidaan luontevalla tavalla liittää päätösvaihtoehtojen paremmuusjärjestyksen määrittämisen yleiseen ja tärkeään ongelmaan. Tämä ongelma on eräänlainen päätäntätieteen "viisasten kivi": sen ratkaiseminen poistaisi päätöksentekijän ehkä suurimman huolen aiheen. Siksi on ongelman ratkaisuyrityksiin suhdauduttava vain alustavina ratkaisuuina, joiden avulla

ongelman monitahoisuuden ja tavanomaisten ratkaisujen välistä epäsuhtaa voidaan pienentää.

Tarkastellaan kysymystä kaavamaisen esimerkin avulla.

Tennispelaajien paremmuus. Olkoon ongelmana valita toimintavaihtoehto äärellisestä vaihtoehtojen joukosta. Havainnollistamme tilannetta olettamalla, että vaihtoehtoja on vain kaksi:

a_1 : valitaan vaihtoehto, jonka tiedetään olevan 10. paras tutkittujen 10 000 satunnaisesti valitun tapauksen joukossa.

a_2 : valitaan vaihtoehto, jonka tiedetään olleen paras tutkittujen 1 000 tapauksen joukossa.

Ajatellaan vaikkapa, että vaihtoehto a_1 vastaa 10. parasta tennispelaajaa (kuinka se sitten saataisiinkin selville) 10 000 pelaajan joukossa ja vaihtoehto a_2 parasta pelaajaa satunnaisesti valitun 1 000 pelaajan joukossa.

Kumpaa vaihtoehtoa pidetään parempana?

Kysymykseen ei tietenkään voida antaa selvää vastausta ellei valinnan seurauksia ja siis valintatilannetta kuvata tarkemmin. Jos tarkoituksena on löytää varma ja tasaisen hyvä pelaaja, joka suurella varmuudella voittaa tavalliset helpot kyläkilpailut, kannattaa valita 1. vaihtoehto.

Jos sen sijaan pelaaja on lähetettävä tenniksen maailmanmestaruuskilpailuihin Wimbledoniin ja ainoastaan etusijojen suoritukseen kiinnitetään huomiota, ei varmastikaan 1. vaihtoehto ole hyvä, vaan kannattaa riskeerata ja valita 2. vaihtoehto.

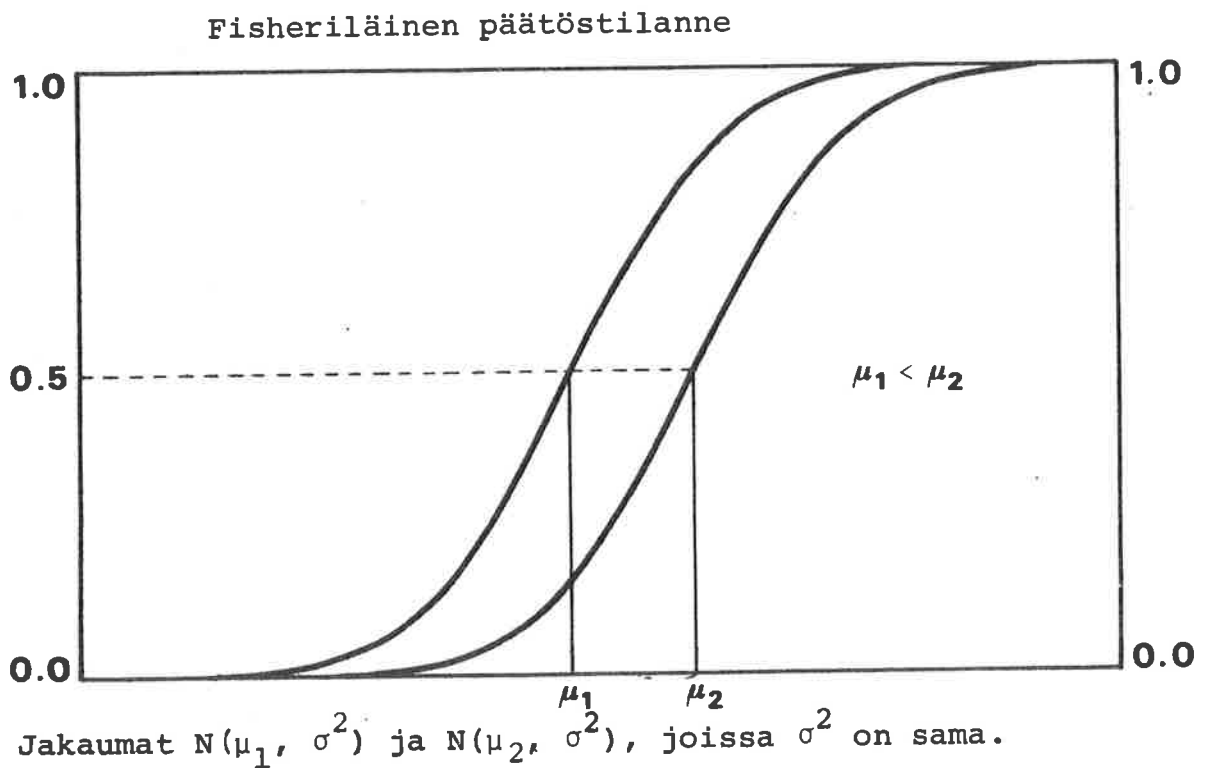
Sillä jos huippupelaajia on paikalla ei 10. paras 10 000:sta varmastikaan voita. Sen sijaan sitä, että pelaaja on paras 1 000:sta pelaajasta, ei voida laskea aivan raskauttavaksi heikkoudeksi.

Yleisemmin: promillemiehenä olemisen pienessä otoksessa ei ole vielä raskas vika, mutta suurten otosten promillemiehiä parempia varmasti löytyy. Ei paraskaan pienessä otoksessa voi olla muuta kuin paras ja maailmanmestariakin on "vain paras" kyläkilpailuissa.

Esimerkiksi yhden havainnon otoksessa paras on myös huonoin eli vain oma itsensä. Tällöin ei ole mitään vertailua suoritettu. Matemaattisesti tätä tilannetta vastaan seuraava tulos: Jos muuttujan x kertymäfunktio $F(x)$ on jatkuva, on muuttujan $F(x)$ jakauma Beta(1,1) eli (0,1)-välin tasajakauma.

Molemmat mainitut pelaajat ovat keskimäärin "yhtä hyviä", mutta 2-pelaaja voi olla paljon parempi tai huonompi kuin 1-pelaaja, jonka ominaisuudet tunnetaan tarkasti. Näin ollen valintaongelma poikkeaa oleellisesti R.A. Fisherin tarkastelusta vaihtoehtojen valintaongelmasta, jossa eri vaihtoehtojen ominaisuuksia kuvaavat normaaliset ja vakiovarianssiset

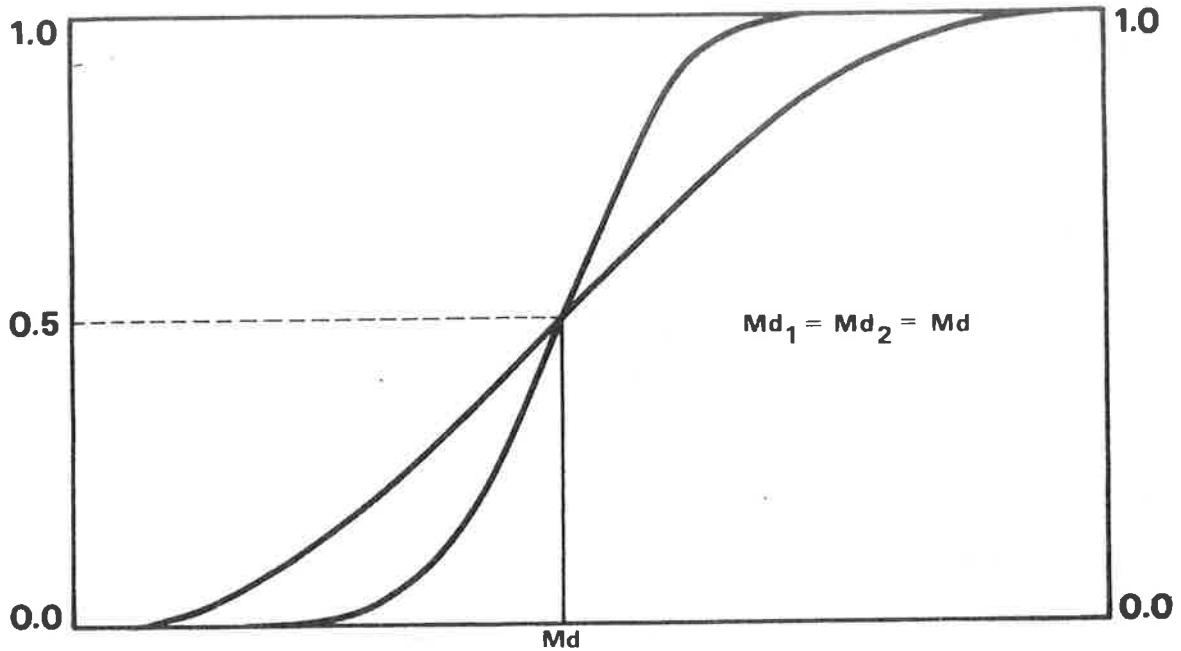
satunnaismuuttujat. Tällöin vaihtoehtojen vertailu suoritetaan varianssianalyysillä. Fisherin tutkimassa päätöstilanteessa ei varsinaisesti eri vaihtoehtojen ominaisuuksia tarvitse punnita, sillä esitetty ratkaisu on kaikissa valintatilanteissa paras menettely. Tämä johtuu siitä, että eri vaihtoehtoihin liittyvät normaalijakaumat saadaan pelkällä siirrolla yhtymään:



Pelkästään vakiovariانسsioletuksen muuttaminen mutkistaa ongelman "ratkaisemattomaksi", kuten esim. Fisher-Behrens-testiin liittyvä keskustelu osoittaa. Tällöin nimittäin eri vaihtoehtoihin liittyvät kertymäfunctiot leikkaavat toisensa, jolloin vaihtoehtojen keskinäisiä etuja ja haittoja on todella punnittava vastakkain.

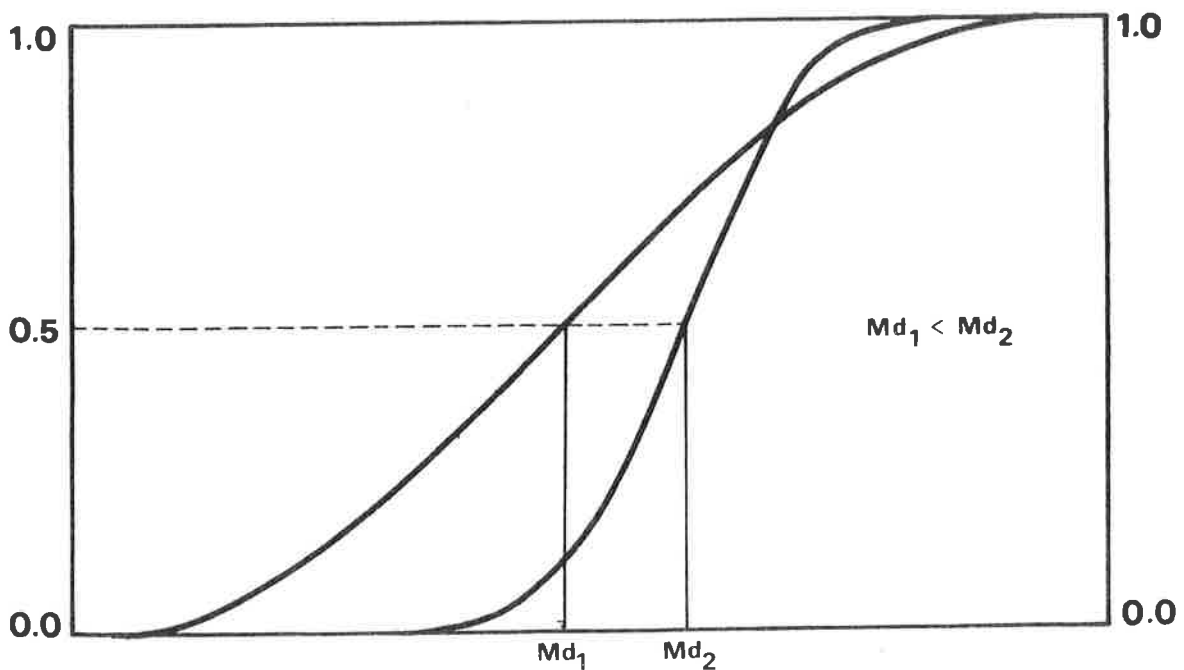
Sen sijaan "tennispelaajien valintaongelmassa" vaihtoehtoihin a_1 ja a_2 liittyvät kertymäfunktiot ovat seuraavaa tyyppiä:

Tennispelaajaongelman päätöstilanne



Jakaumien sijainti on lähes sama, mutta hajonnat ovat eri suuret.

Jos vielä jakaumien sijainnit poikkeavat joudutaan seuraavan kuvion esittämään päätöstilanteeseen.



Näissä päätöstilanteissa päätöksen on perustuttava vaihtoehtoihin liittyvien jakaumien ja valinnan mahdollisten seurauksien välisten yhteyksien analysointiin.

Kuinka hyvä on $k:s$ pelaaja $n:stä$? Kuinka voidaan näin valitun pelaajan hyvyttä (ominaisuuksia) kuvailla?

Jo se, että on mahdollista selvittää edellä esitettyjä asioita matematiikasta tietämättömillekin henkilöille osoittaa, että tarkasteltavana ovat yleisesti käytetyt vertailuperiaatteet.

Peruskäsitteet vertailussa ovat: kuinka suuresta joukosta havainto on saatu, monesko havainto oli suuruusjärjestyksessä ja kuinka tarkasteltava muuttuja jakaantuu perusjoukossa.

Viimeksi mainitun perusjoukon jakauman on koko ajan oletettu olevan kiinteä, mutta tuntematon: itse asiassa vertailun kannalta sen tunteminen ei olekaan tarpeellista. (Vasta käytettäessä eo. kehikkoa deskriptiivisesti, jolloin perusjoukon jakaumasta ei tarvitse olla tietoa tai perusjoukkoa ei asiallisesti ole olemassa, tulee perusjoukon jakauman valinta oleelliseksi kysymykseksi. Eri perusjoukon jakauman valinnoilla saavat käsitteelyt, esim. pelaajat, erilaiset luonnehdinnat.)

Seuraava tulos on matemaattisesti triviaali, mutta tulkinnallisesti hyvin tärkeä:

Olkoon x_p perusjoukon 100p prosenttipiste eli $F(x_p) = p$.
Tällöin todennäköisyys, että k:s havainto n:stä $\underline{x}_k(n)$ on
pienempi kuin x_p (eli voittaa 100p prosentin miehen perus-
joukossa) on

$$(3.7) \quad P(\underline{x}_k(n) \leq x_p) = P(F(\underline{x}_k(n)) \leq F(x_p)) \\ = P(\underline{y}_k(n) \leq p)$$

Jälkimmäinen todennäköisyys voidaan lukea suoraan
Beta(k, n-k+1)-jakauman kertymäfunktioista.

Muunnos $\underline{x}_k(n) \rightarrow \underline{y}_k(n) = F(\underline{x}_k(n))$ muuntaa siis alkuperäisen
x-muuttujan vaihtelualueen väliksi (0,1). Tämän välin pis-
teet ovat tulkittavissa todennäköisyyksiksi tai prosenteik-
si, jotka vastaavat perusjoukon tiettyjä 100p-prosentin ha-
vaintoja. Muunnos on analoginen erityisesti tulojakaumien
tarkastelussa yleisen Lorenz-käyräesityksen kanssa.

Tähän tulokseen perustuu, että on järkevää poistaa kanta-
funktion $F(x)$ vaikutus ja siirtyä eo. muunnoksen avulla yksik-
köneliön tarkasteluihin. Tarkastelemalla yksikköneliön kerty-
mäfunktioita saadaan konkreettinen käsitys n:stä valitusta
havainnosta k:nneksi parhaimman mahdollisuuksista verrattuna
satunnaisesti valittuun perusjoukon havaintoon (toisin sanoen
perusjoukon "todellisiin" prosenttipisteisiin tai prosentti-
miehiin verrattuna).

Seuraavassa esitetään eräitä yleisiä jakauman arvostamista koskevia tuloksia ja sovelletaan niitä betajakaumiin.

Lause 2: Olkoon F ja G kertymäfunktioita. Jos kertymäfunktioon F liittyvä satunnaismuuttuja on kertymäfunktioon G liittyvää muuttujaa stokastisesti pienempi eli

$$\forall x : F(x) \geq G(x)$$

niin kaikilla jatkuvilla, rajoitetuilla ja kasvavilla funktioilla $g : \mathbb{R} \rightarrow \mathbb{R}$

$$\int_{-\infty}^{\infty} g(x) dF(x) \leq \int_{-\infty}^{\infty} g(x) dG(x)$$

ja kääntäen.

Siis jos G on F :n oikealla puolella niin jokaisen kasvavan funktion g odotusarvo jakauman G suhteen on suurempi kuin g :n odotusarvo F :n suhteen. Tämä merkitsee sitä, että kaikki arvostusfunktiot antavat stokastisesti pienemmälle vaihtoehdolle pienemmän arvostustuloksen kuin stokastisesti suuremmalle.

Tod. $\int g dF - \int g dG = \int g d(F-G)$

Koska g on jatkuva ja $(F-G)$:n variaatio on rajoitettu voidaan suorittaa osittaisintegrointi (Katso Rudin (14) s. 122):

$$\int_a^b g d(f-g) = \int_a^b g(F-G) - \int_a^b (F-G)dg$$

$$\rightarrow g(\infty) [1-1] - g(-\infty) 0-0 - \int_{-\infty}^{\infty} (F-G)dg, \text{ kun } a \rightarrow -\infty, b \rightarrow \infty$$

$$= - \int_{-\infty}^{\infty} (F-G)dg \leq 0$$

Viimeinen yhtäsuuruus pätee ainakin jos g on rajoitettu \square

Lauseen soveltamista voidaan useiden jakaumaperheiden tapauksessa yksinkertaistaa valitsemalla g -funktiot jostain em. ehdot täyttävästä funktiojonosta $g_n(x)$ (ns. testifunktiot). Tällöin saadaan seuraavan tyyppisiä lauseita:

Lause 3: Olkoot F ja G jakaumaperheen \mathcal{F} kertymäfunktioita.

Tällöin seuraavat ehdot ovat yhtäpitäviä:

$$1. \quad \forall n : \int g_n dF \leq \int g_n dG$$

$$2. \quad \forall x : F(x) \geq G(x)$$

Jos esimerkiksi \mathcal{F} on betajakaumien $Beta(a,b)$ joukko riittää valita

$$(3.8) \quad g_n(x) = x^n, \quad x \in [0,1]$$

Johdetaan testifunktiojonon (3.8) avulla perustulokset betajakaumien arvostamisesta.

Lause 4: Olkoon $\underline{x} \sim \text{Beta}(a, b)$ ja $v(a, b) = \text{Eg}_n(\underline{x}) = \text{E}\underline{x}^n =$ muuttujan \underline{x}^n odotusarvo a :n ja b :n funktiona. Tällöin

$$1. \quad v(a, b) = \text{E}\underline{x}^n = \frac{\Gamma(a+n) \Gamma(a+b)}{\Gamma(a) \Gamma(a+b+n)}$$

$$2. \quad v(a, b+1) = \frac{\Gamma(a+n) \Gamma(a+b+1)}{\Gamma(a) \Gamma(a+b+n+1)} = v(a, b) \left(\frac{a+b}{a+b+n} \right) \\ \leq v(a, b)$$

$$3. \quad v(a-1, b+1) = \frac{\Gamma(a-1+n) \Gamma(a+b)}{\Gamma(a-1) \Gamma(a+b+n)} = v(a, b) \left(\frac{a-1}{a-1+n} \right) \\ \leq v(a, b), \text{ kun } a > 1$$

$$4. \quad v(a+1, b) = \frac{\Gamma(a+n+1) \Gamma(a+b+1)}{\Gamma(a+1) \Gamma(a+b+n+1)} = v(a, b) \frac{(a+n)(a+b)}{a(a+b+n)} \\ \leq v(a, b)$$

Jos pieniä \underline{x} :n arvoja pidetään suurina parempina (eli arvos-
tustuloksien ja sijalukujen järjestykset vastaavat toisiaan)
voidaan betajakaumien välille määritellä seuraava osittais-
järjestys:

$$\text{Beta}(\hat{a}, \hat{b}) \succeq \text{Beta}(a, b) \text{ joss } \forall x : \hat{F}(x) \geq F(x)$$

Lauseen 4 mukaan pätee

$$2. \quad \text{Beta}(a, b+1) \succeq \text{Beta}(a, b)$$

$$3. \quad \text{Beta}(a-1, b+1) \succeq \text{Beta}(a, b)$$

$$4. \quad \text{Beta}(a,b) \gtrsim \text{Beta}(a+1, b)$$

Kun jakauman $\text{Beta}(a,b)$ parametrit a ja b tulkitaan järjestysluvuiksi laskettuna toisaalta paremmuus- ja huonommuusjärjestyksessä $a+b-1 = n$ suuruisessa otoksessa saadaan seuraavat tulkinnat:

$$2. \quad \begin{aligned} a:s \text{ paras } (n+1):\text{stä} &\gtrsim a:s \text{ paras } n:\text{stä} \\ (b+1):s \text{ huonoin } (n+1):\text{stä} &\gtrsim b:s \text{ huonoin } n:\text{stä} \end{aligned}$$

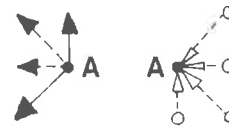
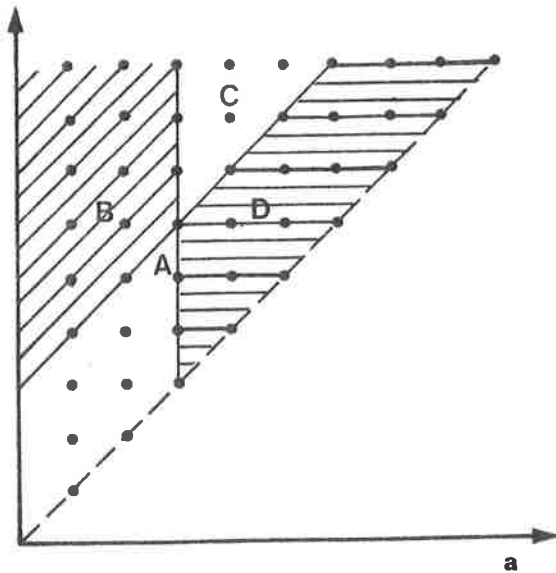
$$3. \quad \begin{aligned} (a-1):s \text{ paras } n:\text{stä} &\gtrsim a:s \text{ paras } n:\text{stä} \\ (b+1):s \text{ huonoin } n:\text{stä} &\gtrsim b:s \text{ huonoin } n:\text{stä} \end{aligned}$$

$$4. \quad \begin{aligned} a:s \text{ paras } n:\text{stä} &\gtrsim (a+1):s \text{ paras } (n+1):\text{stä} \\ b:s \text{ huonoin } n:\text{stä} &\gtrsim b:s \text{ huonoin } (n+1):\text{stä} \end{aligned}$$

Samalla järjestyssijalla olevista havainnoista suuremman otoksen havaintoa pidetään parempana. Saman kokoisten otosten havainnoista paremmalla järjestyssijalla olevaa pidetään parempana. Neljännessä kohdassa intuitiivinen tulkinta saadaan huonommuusjärjestyksen avulla: samalla huonommuussijalla olevista havainnoista suuremman otoksen havainto on huonompi.

Jos kuvataan betajakaumat tason pisteinä siten, että pystyakselillä on "otoskoko" $n = a+b-1$ ja vaaka-akselilla on "sijaluku" a saadaan eo. osittaisjärjestykselle havainnollinen esitys

$$n = a + b - 1$$



Vieressä oleva kuvio muistuttaa, että preferenssi kasvaa pystyriveittäin ylös, vaakariveittäin vasemmalle ja vino-
riveittäin alas ja ylös. Katkoviivoilla osoitetut preferens-
sisuhteet seuraavat jatkuvilla viivoilla osoitetuista prefe-
rensseistä. Näin ollen jakauman A vasemmalla puolella olevan
viivoitetun alueen jakaumia on aina preferoitava A:han näh-
den.

Huomautettakoon, että jokaista pistettä voidaan edellisen
nojalla verrata myöskin neljään muuhun lähipisteeseen, joiden
koordinaatit poikkeavat vertailupisteestä korkeintaan ykkö-
sellä. Nämä lähipisteet jäävät preferenssijärjestyksessä
heikommalle sijalle kuten oheinen kuva osoittaa. Tästä
seuraa, että jakaumaa A on preferoitava sen oikealla puolel-
la olevan viivoitetun alueen jakaumiin nähden.

Vaikkakin lähipisteistä voidaan vertailla toisiinsa riidat-
tomasti, löytyy pisteitä, joiden kesken ei preferenssiä voida
näin määrätä. Esimerkiksi jakaumien $Beta(a, b)$ ja $Beta(a+1, b+1)$

vertailu ei ole riidattomasti suoritettavissa, koska niiden kertymäfunctiot leikkaavat; jakaumien edut on siis punnittava vastakkain.

Näin määräytyvä epämääräisyysalue on kuviossa eräs viivottoman alueen osajoukko. Tämän osajoukon jakaumien ja A:n välinen järjestys on päätöksentekijän määrättävissä ja perustuu siis päätöksentekijän harkintaan. Tilanne on vastaava kuin Einsteinin erikoisessa suhteellisuusteoriassa, jossa ajallisen epämääräisyysalueen pisteiden järjestys määräytyy tarkkailijan liiketilan perusteella. (Katso esim. Nevanlinna (15) ss. 142-151 tai Laurikainen (16) ss. 45-47.)

4. BETAJAKAUMAN ML-YHTÄLÖIDEN RATKAISEMINEN ELI
BIJEKTION $(a,b) \rightarrow (f,g)$ KÄÄNTEISFUNKTION
APPROKSIMOINTI

4.1. YLEISTÄ

Betajakauman parametrien ml-estimointi on tiivistynyt yhtälöpariin

$$(4.1.1) \quad \psi(a+b) - \psi(a) = u$$

$$(4.1.2) \quad \psi(a+b) - \psi(b) = v,$$

jossa u ja v riippuvat geometrisista keskiarvoista
 $f = \text{Geom}(\underline{x})$ ja $g = \text{Geom}(1-\underline{x})$ seuraavasti

$$(4.1.3) \quad u = E \log\left(\frac{1}{\underline{x}}\right) = \log\left(\frac{1}{f}\right) \text{ eli } f = e^{-u}$$

$$(4.1.4) \quad v = E \log\left(\frac{1}{1-\underline{x}}\right) = \log\left(\frac{1}{g}\right) \text{ eli } g = e^{-v}.$$

Yhtälöt (4.1.1 - 2) määrittelevät kääntäen yksikäsitteisen kuvauksen joukolta $B = \{(a,b) \mid a > 0, b > 0\}$ joukolle $C = \{(u,v) \mid u > 0, v > 0, e^{-u} + e^{-v} < 1\}$. Tätä kuvausta on tosin vaikea havainnollistaa, koska joukot B ja C eivät ole rajoitettuja eikä tunnusluvulla (a,b) ja (u,v) ole suoraa tilastollista tulkintaa. Tämän vuoksi tarkastelemme usein yhtälöparin

$$(4.1.5) \quad \psi\left(\frac{1-\theta}{\theta}\right) - \psi\left(p\left(\frac{1-\theta}{\theta}\right)\right) = \log\left(\frac{1}{f}\right)$$

$$(4.1.6) \quad \psi\left(\frac{1-\theta}{\theta}\right) - \psi\left(q\left(\frac{1-\theta}{\theta}\right)\right) = \log\left(\frac{1}{g}\right), \quad q = 1-p$$

määrittelemää bijektiota yksikköneliöltä

$$A = \{(p, \theta) \mid 0 < p < 1, 0 < \theta < 1\} \text{ kolmiolle}$$

$$B = \{(f, g) \mid 0 < f < 1, 0 < g < 1, f+g < 1\}.$$

Tämä bijektio muodostuu kolmen bijektion $\varphi_1, \varphi_2, \varphi_3$ kompositiona $\varphi_3 \circ \varphi_2 \circ \varphi_1$, kun bijektiot määritellään seuraavasti:

$$(4.1.7) \quad \varphi_1 : A \rightarrow B, \text{ jolle}$$

$$(p, \theta) \xrightarrow{\varphi_1} (a, b) = \left(p\left(\frac{1-\theta}{\theta}\right), q\left(\frac{1-\theta}{\theta}\right)\right)$$

$$(4.1.8) \quad \varphi_2 : B \rightarrow C, \text{ jolle}$$

$$(a, b) \xrightarrow{\varphi_2} (u, v) = (\psi(a+b) - \psi(a), \psi(a+b) - \psi(b))$$

$$(4.1.9) \quad \varphi_3 : C \rightarrow D, \text{ jolle}$$

$$(u, v) \xrightarrow{\varphi_3} (f, g) = (e^{-u}, e^{-v})$$

Bijektiot φ_1 ja φ_3 ovat helposti käännettävissä, joten voimme φ_2 :n ominaisuuksia selvittäessä siirtyä muuttujista (a,b) haluttaessa muuttujiin (p,θ) tai joihinkin niiden yksinkertaisiin muunnoksiin. Samoin voidaan u :n ja v :n sijaan käyttää f :ää ja g :tä.

Ongelman hahmottamisen kannalta on edullista yrittää ensin kääntää suoraan bijektiota $\varphi_3 \circ \varphi_2 \circ \varphi_1$ (kiinnittämättä huomiota sen kompositioluonteeseen) eli lausua p ja θ suoraan muuttujista f ja g . Näin saadut ensimmäiset arviot

$$(4.1.10) \quad p \approx \frac{1-g}{2-f-g}$$

$$(4.1.11) \quad \left(\frac{\theta}{1-\theta}\right) = \frac{1}{a+b} \approx 2(1-f-g)$$

toimivat hyvin, kun $f+g \approx 1$ ja $f \approx g$, jolloin $\theta \approx 0$ ja $p \approx \frac{1}{2}$.

Toisaalta alueella $\theta \approx 1$ kannattaa tarkastella kuvausta $\varphi_2 \circ \varphi_1$ parilta (p,θ) parille (u,v) . Tämän käänteiskuvaus on likimain muotoa

$$(4.1.12) \quad p \approx \frac{1}{1 + \sqrt{\frac{u}{v}}}$$

$$(4.1.13) \quad \left(\frac{\theta}{1-\theta}\right) = \frac{1}{a+b} \approx \sqrt{uv}$$

Kuitenkin varsinaiset tarkkuuskaavat kehitetään kuvauksien $(a,b) \xrightarrow{\varphi_2} (u,v)$ tai $(a,b) \xrightarrow{\varphi_3 \circ \varphi_2} (f,g)$ käänteiskuvauksille.

Kun (a,b) on lausuttu joko (u,v) :n tai (f,g) :n funktiona, saadaan a :n ja b :n lausekkeet kuten $p = \frac{a}{a+b}$, $\theta = \frac{1}{a+b+1}$, $x = a+b$ jne. sijoituksella.

Vaikka tutkimuksen päätarkoituksena on approksimoida parametreja a ja b yksinkertaisilla geometrinen keskiarvojen f ja g funktioilla on liitteessä 3 esitetty Newtonin iterointimenetelmään perustuva algoritmi, joka laskee a :n ja b :n arvot, kun u ja v ovat annettuja. Tätä algoritmia voidaan käyttää haluttaessa otossuureiden u ja v havaintoarvoja vastaavat maximum likelihood-estimaatit a ja b . Liitteenä oleva tietokoneohjelma on laadittu GE 600 BASIC-kielelle.

Liitteenä 4 olevan taulukon avulla on laadittu sivuilla 56-57 esitetyt diagrammat I ja II, joiden avulla saa havainnollisen käsityksen kuvauksien $(f,g) \rightarrow (p,\theta)$ ja $(f,g) \rightarrow (a,b)$ ominaisuuksista.

Eräs huomionarvoinen diagrammoista ilmenevä seikka on, että geometrinen keskiarvojen f ja g summa on korkeintaan yksi, kuten aiemmin todistettiin. Näin ollen kuvaukset $(f,g) \rightarrow (p,\theta)$ ja $(f,g) \rightarrow (a,b)$ ovat määritellyt vain kolmiossa $D = \{(f,g) | f > 0, g > 0, f+g < 1\}$, jonka reunoilla kuvaukset tulevat epämääräisiksi.

Diagonaalilla $f+g = 1$ vastaava betajakauma on kokonaan keskittynyt odotusarvossa p kohdalle. Tällöin kaavojen (4.1.10 - 11) mukaan $p = 1-g = f$ ja $\theta = 0$. Mikäli a ja b molemmat ovat suuria on $Beta(a,b)$ likimain normaalijakauman muotoinen, jolloin $f:n$ ja $g:n$ summa on likimain yksi ja arviot (4.1.10) sekä (4.1.11) toimivat hyvin. Tätä tilannetta tarkastellaan luvussa 4.2.

Diagramma I : Parametrien θ ja p nivookäyriä geometrinen keskiarvojen f ja g määräämässä koordinaatistossa

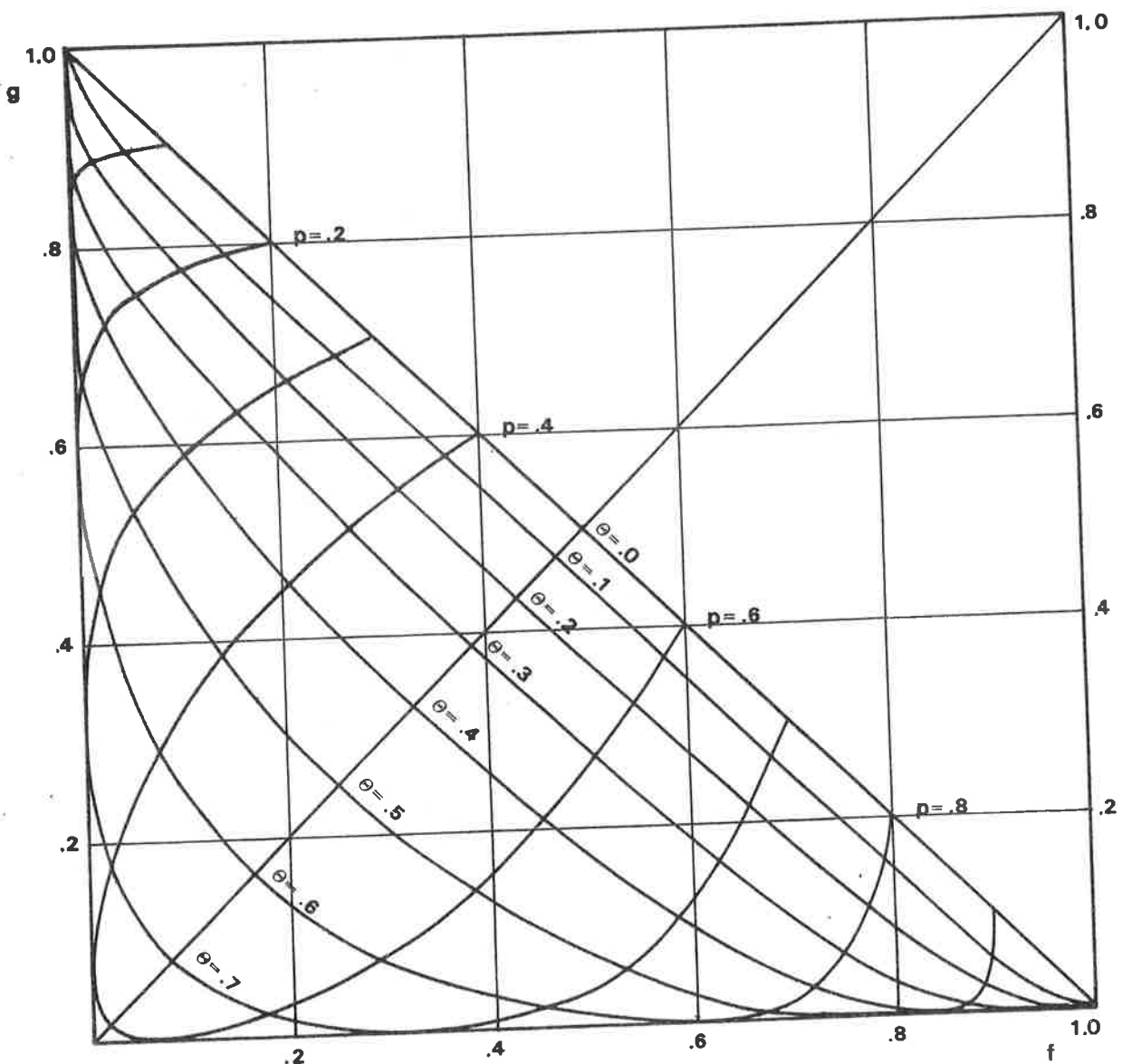
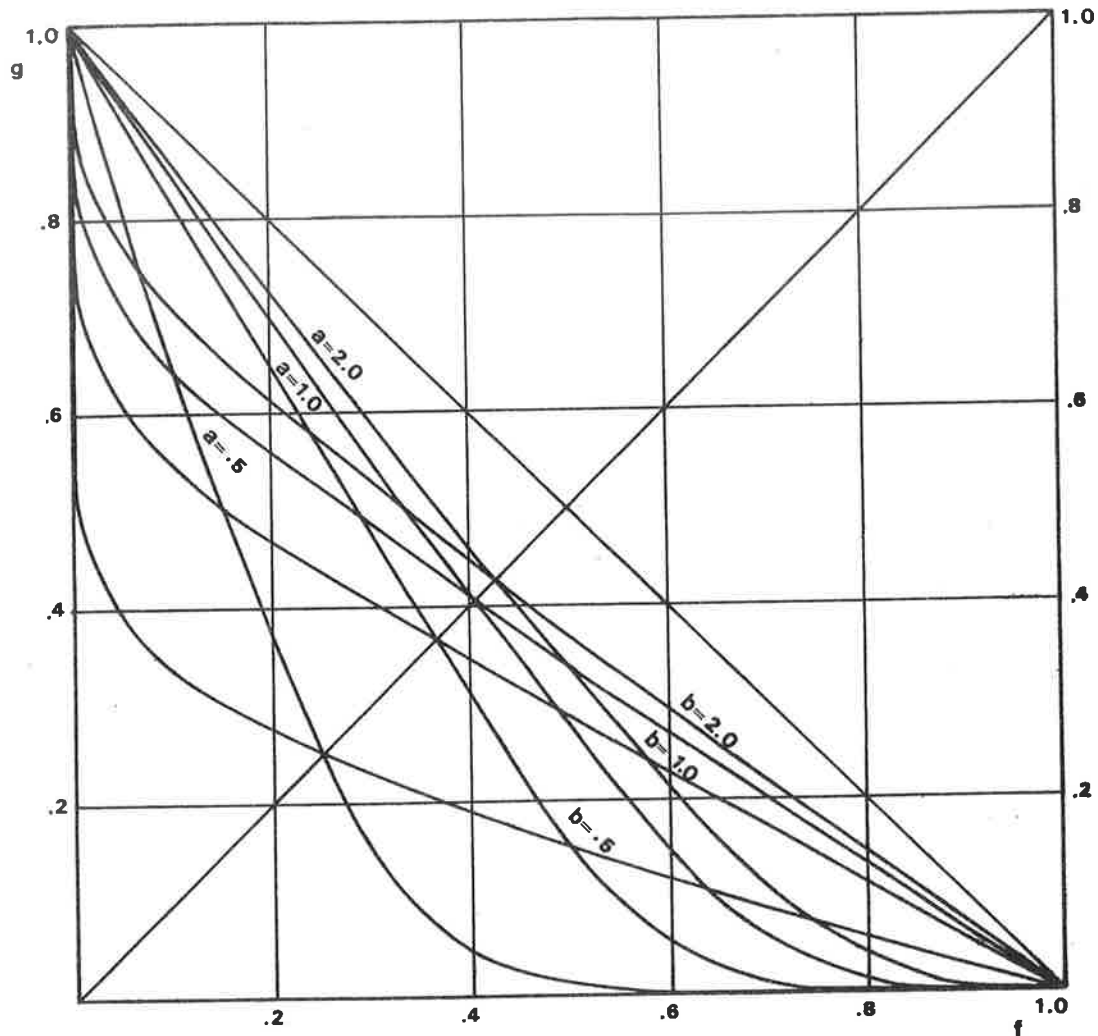


Diagramma II: Parametrien a ja b nivookäyriä geometrinen keskiarvojen f ja g määräämässä koordinaatistossa

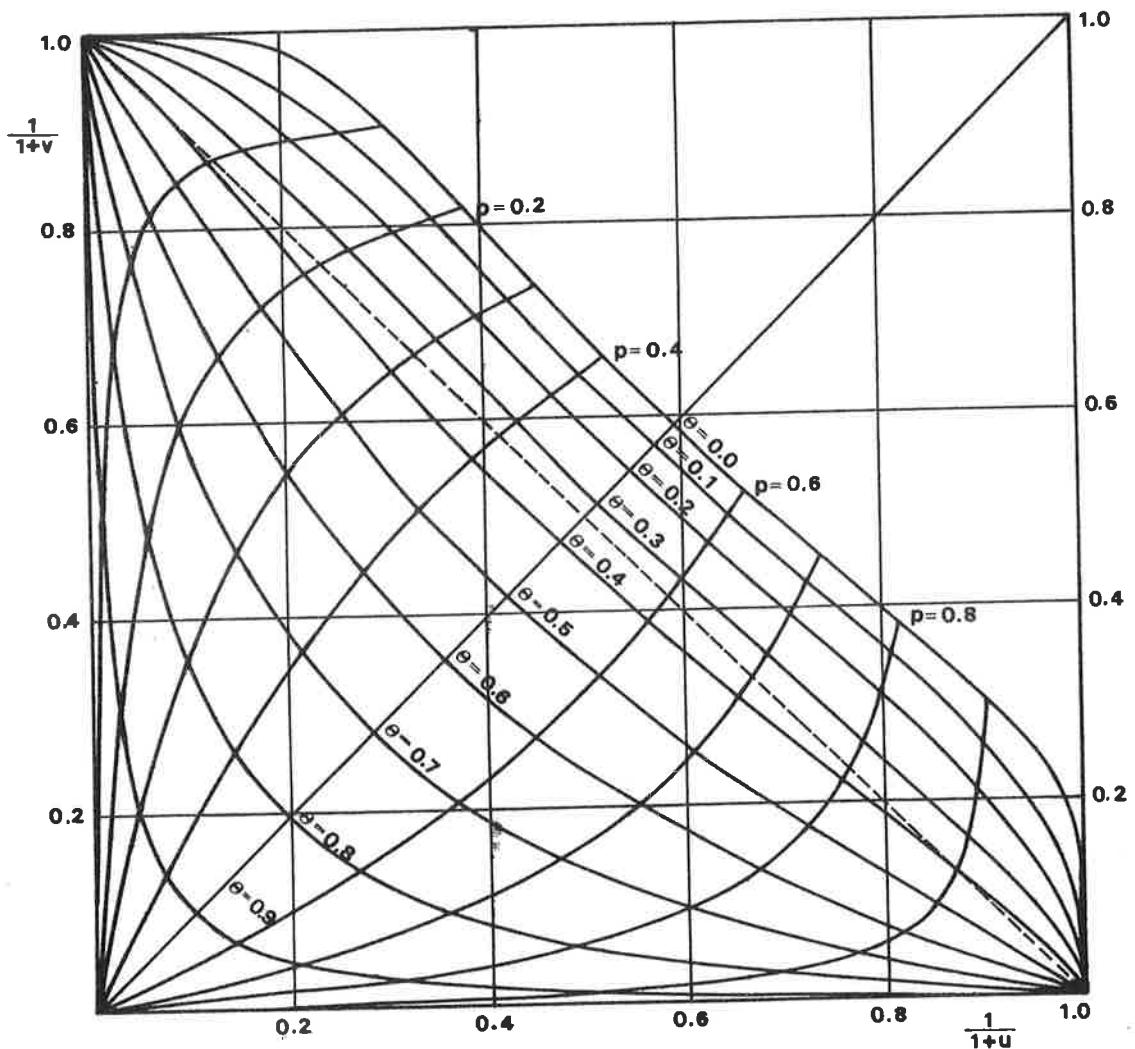


Mikäli geometrisista keskiarvoista $f = e^{-u}$ ja $g = e^{-v}$ toinen on lähellä nollaa (eli u tai v on suuri), ollaan diagrammoissa lähellä toista koordinaattiakselia. Tällöin a ja b molemmat ovat pieniä¹⁾, joten $Beta(a,b)$ on U :n muotoinen. Tässä tapauksessa arviot ovat tyyppiä (4.1.12) ja

1) Paitsi jos $f+g \approx 1$, jolloin ollaan toisen kolmion nurkkapisteen lähistöllä. Tätä tapausta tarkastellaan seuraavaksi.

(4.1.13). Koska p ja θ sekä a ja b muuttuvat lähellä koordinaattiakseleita hyvin nopeasti, eivät diagrammat I ja II anna tarkkaa kuvaa tilanteesta. Diagramma III on diagrammaa I vastaava kuvio, jossa $f:n$ ja $g:n$ sijaan on käytetty muuttujina $(0,1)$ -välin suureita $1/(1+u)$ ja $1/(1+v)$. Tällöin $p:n$ ja $\theta:n$ nivookäyrät saadaan selvemmin esiin. U-muotoisten beta-jakaumien tapausta tarkastellaan luvussa 4.3.

Diagramma III: Parametrien p ja θ nivookäyriä suureiden $1/(1+u)$ ja $1/(1+v)$ määräämässä koordinaatistossa



Edellä käsitellyissä tapauksissa laajennettiin likelihood yhtälöiden ratkaisut singulaarisista tapauksista $f+g = 1$ ja $f = 0$ tai $g = 0$ lähtien. Tämä merkitsee (diagrammien avulla ilmaistuna) kolmion reunoilla tarkkojen kaavojen yleistämistä kohti kolmion keskustaa.

Reunalla $f+g = 1$ on estimoitujen betajakaumien todennäköisyysmassa keskittynyt kokonaan painopisteeseensä. Mikäli ollaan f -akselilla, on $\theta = 1$ ja $p = 1$, mikä merkitsee estimoidun betajakauman keskittyneen x -akselin pisteeseen 1 ja yhtyvän singulaariseen binomijakaumaan $\text{Bin}(n = 1, p = 1)$. Vastaavasti g -akselin tapauksessa estimoitu betajakauma yhtyy binomijakaumaan $\text{Bin}(n = 1, p = 0)$. Kuten diagrammasta I havaitaan, ovat kolmion kaikki nurkkapisteet erikoispisteitä $p:n$ tai $\theta:n$ kannalta. Nurkissa $(f,g) = (0,1)$ ja $(f,g) = (1,0)$ on θ epämääräinen. Tämä voidaan tulkita siten, että betajakauma voi keskittyä esimerkiksi x -akselin pisteeseen 0 äärettömän monella tavalla: normaalijakauman, gammajakauman tai $U:n$ muotoisten betajakaumien välityksellä, kun näiden painopiste lähenee origoa.

Vastaavasti piste $(f,g) = (0,0)$ on erikoispiste $p:n$ suhteen, samalla kun $\theta = 1$. Kyseinen nurkkapiste sisältää kaikki singulaariset $U:n$ muotoiset betajakaumat eli binomijakaumat $\text{Bin}(n = 1, p)$.

Lukijan on hyödyllistä havainnollistaa itselleen kahden muuttujan funktioiden $\theta = \theta(f,g)$ ja $p = p(f,g)$ käyttäytymistä diagramman I nivookäyrien avulla, jolloin likelihood yhtälöiden transsendenttinen luonne selvästi ilmenee.

Edellisen nojalla hallitaan singulaarisen pisteen $(f,g) = (0,0)$ ympäristö kaavojen (4.1.12 - 13) avulla. Sen sijaan nurkkapistet $(f,g) = (0,1)$ ja $(f,g) = (1,0)$ ovat huomattavasti vaikeammin käsiteltävissä. Yksinkertaiset ratkaisut, jotka ovat hyviä normaalityyppisille tai U-muotoisille jakaumille eivät sellaisenaan toimi näissä nurkkapisteissä, joihin liittyy lähinnä J-muotoisia betajakaumia. Näitä gammajakauman tyyppisiä betajakaumia tarkastellaan luvussa 4.4.

4.2. NORMAALIJAKAUMAN MUOTOISTEN BETAJAKAUMIEN

TAPAUS: $a \geq 2$ JA $b \geq 2$ SEKÄ $a/b \approx 1$

Tässä luvussa tarkastellaan likelihood yhtälöiden ratkaisemista, kun a ja b molemmat ovat suuria. Tällöin $Beta(a,b)$ on normaalijakauman muotoinen yksihuippuinen jakauma, jossa $f+g \approx 1$. Diagrammassa I siis ollaan lähellä kolmion reunaa $f+g = 1$. Etäisyyttä tästä reunasta mitataan yksinkertaisesti positiivisella suureella

$$(4.2.1) \quad d = \frac{1-f-g}{2} ,$$

jonka avulla edellisen luvun kaavat (4.1.10 - 11) voidaan esittää muodossa

$$(4.2.2) \quad p \approx \frac{1-g}{2-f-g} = \frac{f+d}{1+d}$$

$$(4.2.3) \quad q \approx \frac{1-f}{2-f-g} = \frac{g+d}{1+d}$$

$$(4.2.4) \quad x = a+b \approx \frac{1}{2d}$$

Näiden arvioiden tarkkuus ilmenee liitteestä 6 . Huomattakoon, että $p:n$ ja $q:n$ arviot ovat täysin symmetrisiä ja niiden summa

$$(4.2.5) \quad \frac{f+d}{1+d} + \frac{g+d}{1+d} = \frac{2-f-g}{2-f-g} = 1 ,$$

kuten pitääkin. Yhtälön (4.2.4) mukaan $a+b$ on likimain kääntäen verrannollinen diagonaalista laskettuun etäisyyteen $d = 1-f-g$.

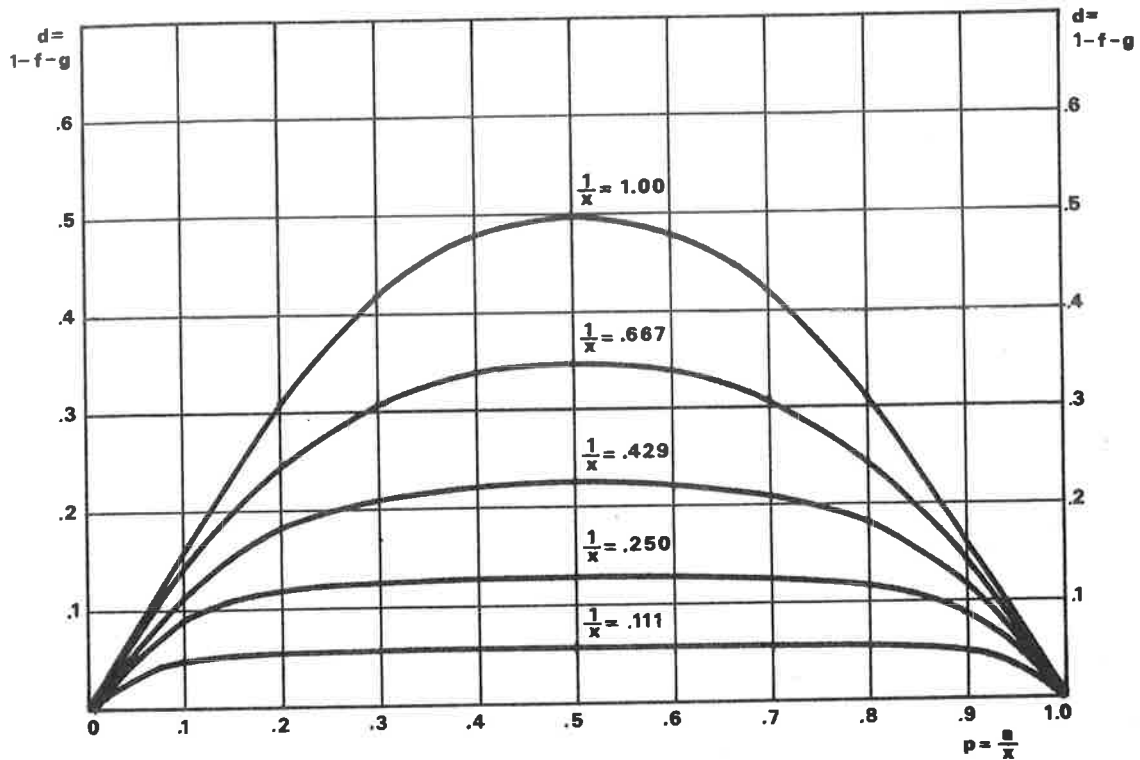
Vastaavat arviot a :lle ja b :lle ovat

$$(4.2.6) \quad a = px \approx \left(\frac{f+d}{1+d}\right) \frac{1}{2d}$$

$$(4.2.7) \quad b = qx \approx \left(\frac{g+d}{1+d}\right) \frac{1}{2d} .$$

Oleellista approksimaatioissa on, että suurilla a :n ja b :n arvoilla $x = a+b$ on likimain vain etäisyyden $d = 1-f-g$ funktio, joten 'vinouden' vaikutus on hyvin pieni. Vasta kun f ja g ovat aivan eri suuruusluokkaa alkavat arviot (4.2.2 - 4) pettää ja f :n ja g :n yksilölliset vaikutukset on otettava huomioon. Diagrammassa IV olevat käyrät ovat $\frac{1}{x}$:n nivookäyriä p :n ja d :n määräämässä koordinaatistossa. Kun d pienenee, pienenee (jokaisella kiinteällä p :llä) myös suure $\frac{1}{x}$. Samalla $\frac{1}{x}$:n nivookäyrät suoristuvat, mikä merkitsee, että $\frac{1}{x}$ on yhä tarkemmin vain d :n funktio. Jos $\frac{1}{x}$:n nivookäyrät olisivat suoria, olisi $\frac{1}{x}$ vain d :n funktio.

Diagramma IV: Suureen $1/x$ nivookäyrät p :n ja d :n määräämässä koordinaatistossa



Voimmekin todeta, että symmetrisessä tapauksessa $p = \frac{1}{2}$ (eli $a = b$) tarkat $\frac{1}{x}$:n arviot toimivat hyvin myös, kun $p \approx \frac{1}{2}$ erityisesti pienillä d :n arvoilla. Tämän tuloksen avulla saadaan symmetrisen tapauksen avulla johdettua yleisemminkin käyttökelpoisia tuloksia. Johdetaan ensin x :n arvio (4.2.4) käyttämällä m_1 -yhtälöiden esitysmuotoa

$$(4.2.8) \quad x = a+b = \psi^{-1}(\psi(x)-u) + \psi^{-1}(\psi(x)-u).$$

Koska suurilla y :n arvoilla

$$(4.2.9) \quad \psi^{-1}(y) \approx e^y + \frac{1}{2},$$

saadaan esimerkiksi

$$(4.2.10) \quad \psi^{-1}(\psi(x)-u) \approx e^{\psi(x)-u+\frac{1}{2}} = f e^{\psi(x)+\frac{1}{2}}.$$

Myöhemmin johdettavissa x :n tarkkuuskaavoissa ratkaistaan (4.2.8):aa vastaava yhtälö ensin x :n funktion $e^{\psi(x)}$ suhteen ja tämä edelleen x :n suhteen. Tässä sijoitetaan sen sijaan

$$(4.2.11) \quad e^{\psi(x)} \approx x - \frac{1}{2},$$

jolloin saadaan seuraavat yhtäpitävät arviot:

$$(4.2.12) \quad x \approx f(x - \frac{1}{2}) + \frac{1}{2} + g(x - \frac{1}{2}) + \frac{1}{2}$$

$$(4.2.13) \quad (x - \frac{1}{2}) \approx (f+g)(x - \frac{1}{2}) + \frac{1}{2}$$

$$(4.2.14) \quad d(x - \frac{1}{2}) \approx \frac{1}{2}, \quad d = 1-f-g$$

$$(4.2.15) \quad x \approx \frac{1}{2d} + \frac{1}{2}$$

Arviossa (4.2.15) kuitenkin luku $\frac{1}{2}$ on "virheellinen": pienillä d :n arvoilla $\frac{1}{2}$ tulee korvata $\frac{1}{4}$:llä kuten liitteessä 2 symmetrisen tapauksen $a = b$ avulla osoitetaan. Joka tapauksessa x on likimain kääntäen verrannollinen d :n kanssa, kun a ja b ovat suuria.

Odotusarvon $p = \frac{a}{x}$ arvioon (4.2.2) päästään esimerkiksi yhtälöstä

$$(4.2.16) \quad f-g = \frac{e^{\psi(a)} - e^{\psi(b)}}{e^{\psi(x)}}$$

käyttämällä (4.2.10):tä vastaavia arvioita:

$$(4.2.17) \quad f-g \approx \frac{(a - \frac{1}{2}) - (b - \frac{1}{2})}{x - \frac{1}{2}}$$

$$(4.2.18) \quad (f-g)(x - \frac{1}{2}) \approx a-b$$

$$(4.2.19) \quad (f-g)(1 - \frac{1}{2}x) \approx p-q .$$

Käyttämällä vasemmalla puolella geometrisen sarjan summa-kaavaa ja arviota $x \approx \frac{1}{2d}$ saadaan

$$(4.2.20) \quad \frac{f-g}{1+d} \approx p-q ,$$

joka on yhtäpitävä $p:n$ arvion (4.2.2) kanssa. Parametrin x arvioon (4.2.15) päästään myös lähtemällä yhtälöä (4.2.16) vastaavasta $(f+g):n$ yhtälöstä.

Liitteessä 2 symmetrisen tapauksen avulla johdettu $x:n$ arvio (19) eli

$$(4.2.21) \quad x \approx \frac{1-d^2}{2d-d^2}$$

on huomattava parannus arvioon (4.2.15) nähden. Kombinoimalla tämä yksinkertaisiin p:n ja q:n arvioihin, saadaan kaavat

$$(4.2.22) \quad a \approx \left(\frac{f+d}{1+d}\right) \left(\frac{1-d^2}{2d-d^2}\right) = \frac{(f+d)(1-d)}{d(2-d)}$$

$$(4.2.23) \quad b \approx \left(\frac{g+d}{1+d}\right) \left(\frac{1-d^2}{2d-d^2}\right) = \frac{(g+d)(1-d)}{d(2-d)} .$$

Arvioiden (4.2.21-23) tarkkuudet ilmenevät seuraavasta taulukosta.

x	p	$\frac{1-d^2}{2d-d^2}$	$\frac{(f+d)(1-d)}{d(2-d)}$	$\frac{(g+d)(1-d)}{d(2-d)}$
9.0	0.1	10.134	1.020	9.114
	0.2	9.297	1.863	7.433
	0.3	9.100	2.732	6.367
	0.4	9.029	3.612	5.416
	0.5	9.009	4.505	4.505
4.0	0.1	5.679	0.590	5.088
	0.2	4.427	0.898	3.529
	0.3	4.144	1.251	2.893
	0.4	4.046	1.622	2.424
	0.5	4.020	2.010	2.010
2.333	0.1	4.425	0.491	3.934
	0.2	2.885	0.604	2.281
	0.3	2.521	0.772	1.749
	0.4	2.394	0.965	1.429
	0.5	2.361	1.180	1.180
1.5	0.1	3.797	0.466	3.331
	0.2	2.153	0.479	1.674
	0.3	1.722	0.544	1.178
	0.4	1.566	0.639	0.927
	0.5	1.524	0.762	0.762
1.000	0.1	3.256	0.456	2.800
	0.2	1.711	0.419	1.292
	0.3	1.239	0.414	0.826
	0.4	1.052	0.439	0.613
	0.5	1.000	0.500	0.500

Vertailemalla taulukkoa liitteen 4 parametrien a ja b arvoihin havaitaan arvioiden (4.2.22-23) virheiden olevan korkeintaan 2 %, kun $a = b \geq 0.5$. Johtuen pääasiassa $x:n$ arvion (4.2.21) heikkenemisestä kasvavat $a:n$ ja $b:n$ arvioiden virheet, kun p poikkeaa 0.5:stä. Tämä taipumus kuitenkin häviää $x:n$ kasvaessa. Esimerkiksi kun $x = 9$, $p \geq 0.3$ ja $q \geq 0.3$ ovat virheet alle 2 %. Suuremmilla $x:n$ arvoilla, jotka ovat tavanomaisissa sovellutuksissa odotettavasti yleisempiä, ovat arvioiden käyttöalueet vielä laajempia.

Siirrymme varsinaisten tarkkuuskaavojen johtamiseen. Suurilla $a:n$ ja $b:n$ arvoilla yksinkertainen ratkaisumenetelmä perustuu käänteisfunktio- μ otoon (2.4.25). Arvioiden yleispiirteet saadaan esitettyä käyttämällä ψ in käänteisfunktion approksimaatiota

$$(4.2.24) \quad x = \psi^{-1}(y) \approx z + \frac{1}{2} - \frac{1}{24z}, \quad \text{jossa } z = e^y$$

Approksimaation virhe on pienempi kuin $3 \cdot 10^{-2}$, kun $y \geq -1$ eli $x \geq 0.78$ ja virhe häviää nopeasti $y:n$ (ja $x:n$) kasvaessa.

Sijoittamalla ml-yhtälöiden käänteisfunktio- μ otoon

$$(4.2.25) \quad \psi^{-1}(y) = \psi^{-1}(y-u) + \psi^{-1}(y-v)$$

arvio (4.2.24) saadaan

$$(4.2.26) \quad z + \frac{1}{2} - \frac{1}{24z} \approx fz + \frac{1}{2} - \frac{1}{24fz} + gz + \frac{1}{2} - \frac{1}{24gz},$$

joka sievenee muotoon

$$(4.2.27) \quad 2dz \approx 1 - \frac{1}{12z} \left(\frac{1}{f} + \frac{1}{g} - 1 \right), \text{ jossa}$$

$$d = 1-f-g \text{ ja } z = e^Y = e^{\psi(x)}.$$

Seuraavat tulokset voidaan päätellä välittömästi:

- A. Jos $d \approx 0$ ja f ja g ovat selvästi nolasta eroavia, on $2dz \approx 1$ eli $z \approx 1/2d$. Tämä on tilanne, kun a ja b ovat molemmat suuria.
- B. Jos f tai g lähenee nolaa (eli $fg \rightarrow 0$), antaa (4.2.27) täysin virheellisiä arvioita z :lle. Tällöin on a tai b pienempi kuin 0.78 ja käytetty psiin käänteisfunktion approksimaatio (4.2.24) ei toimi. Arviota (4.2.27) on siis korjattava, kun $fg \rightarrow 0$.

Ratkaisemalla arvio (4.2.27) kvadraattisesti muuttujan z suhteen saadaan

$$(4.2.28) \quad z \approx \frac{1}{4d} \left(1 + \sqrt{1 - \frac{2d}{3} \left(\frac{1}{f} + \frac{1}{g} - 1 \right)} \right) = df z_1$$

Tämän jälkeen voidaan a ja b arvioida kaavoilla

$$(4.2.29) \quad a = \psi^{-1}(y-u) \approx fz_1 + \frac{1}{2} - \frac{1}{24fz_1} = df a_1$$

$$(4.2.30) \quad b = \psi^{-1}(y-v) \approx gz_1 + \frac{1}{2} - \frac{1}{24gz_1} = \text{df } b_1$$

Tarkastuksen vuoksi voidaan x arvioida paitsi summana

$a_1 + b_1$ myös suoraan z_1 :n funktiona

$$(4.2.31) \quad x = \psi^{-1}(y) \approx z_1 + \frac{1}{2} - \frac{1}{24z_1} = \text{df } x_1$$

Seuraavasta taulukosta ilmenee arvioiden tarkkuus niiden sovellutusalueella.

x	p	z_1	x_1	a_1	b_1
9.0	0.1	8.0325	8.5273	0.8501	7.6772
	.2	8.4739	8.9689	1.7935	7.1755
	.3	8.4974	8.9925	2.6976	6.2948
	.4	8.5015	8.9967	3.5987	5.3981
	.5	8.5025	8.9976	4.4988	4.4988
4.0	0.1	-	-	-	-
	.2	3.1673	3.6541	0.7223	2.9318
	.3	3.4485	3.9364	1.1788	2.7576
	.4	3.4871	3.9752	1.5894	2.3858
	.5	3.4942	3.9822	1.9911	1.9911
2.333	0.1	-	-	-	-
	.2	-	-	-	-
	.3	1.3805	1.8503	0.5300	1.3203
	.4	1.7352	2.2112	0.8789	1.3323
	.5	1.7742	2.2508	1.1254	1.1254

Havaitaan, että arviot ovat hyvin tarkkoja, kun a ja b ovat molemmat vähintään kaksi. Virhe on tällöin alle 0.5 %. Jos vähintään toinen parametreista a ja b on alle kahden, kasvaa virhe nopeasti, kuten käytettyjen approksimaatiokaavojen mukaan oli odotettavissakin. Lopulta kaava antaa mielettömiä

tuloksia juurrettavan muuttuessa negatiiviseksi (taulukossa olevat viivat osoittavat näitä tapauksia).

Tarkastelemalla yksityiskohtaisemmin arvioita z_1 , x_1 , a_1 ja b_1 havaitaan suhteista (z_1/z) , (x_1/x) , (a_1/a) , (b_1/b) , että virhelähde sisältyy z :n arviioon. Tämän jälkeen sovelletut arviointimenetelmät vain säilyttävät z_1 :n suhteellisen virheen. Tätä varmentaa myös se, että $a_1 + b_1$ on hyvin lähellä x_1 :tä, joka on laskettu suoraan z_1 :n funktiona.

Arvioita z_1 , a_1 ja b_1 voidaan tarkentaa ja niiden sovellutus- aluetta laajentaa seuraavasti. Käytetään psiin käänteisfunktion arvion (4.2.24) sijalla arviota

$$(4.2.32) \quad x = \psi^{-1}(y) \approx z + \frac{1}{2} - \frac{1}{24z + \frac{2}{z+1}}, \text{ jossa } z = e^y.$$

Kun jatketaan vastaavalla tavalla kuin arviota (4.2.27) johdettaessa, saadaan sille seuraava tarkennus:

$$(4.2.33) \quad 2dz \approx 1 - \frac{1}{12z} \left(\frac{1}{f + \frac{1}{12z(fz+1)}} + \frac{1}{g + \frac{1}{12z(gz+1)}} - \frac{1}{1 + \frac{1}{12z(z+1)}} \right)$$

"Ratkaisemalla" 2. asteen yhtälö kuten arviota (4.2.28) johdettaessa päädytään z :n approksimaatioon z_2

$$(4.2.34) \quad z_2 = \frac{1}{4d} \left[1 + \sqrt{1 - \frac{2d}{3}(A(f) + A(g) - A(1))} \right], \text{ jossa}$$

$$(4.2.35) \quad A(v) = \frac{1}{v + \frac{1}{12z_0(vz_0 + 1)}} \quad \text{ja}$$

$$(4.2.36) \quad z_0 = \frac{fg + d^2}{fg \cdot 2d + d^2 \left| \exp\left(\frac{uv}{\sqrt{uv+1.2}}\right) - 1 \right|}$$

Arvio z_0 voidaan tulkita z :n arvioiden $1/2d$ ja

$1/\exp\left(\frac{uv}{\sqrt{uv+1.2}}\right) - 1$ painotetuksi harmoniseksi keskiarvoksi painojen ollessa fg ja d^2 . Tätä voidaan pitää globaalisen z :n arviona, kuten liitteestä 8 havaitaan. Kun a ja b ovat vähintään 2, on z_2 :n suhteellinen virhe pienempi kuin 0.5 % ja virhe häviää nopeasti a :n ja b :n kasvaessa.

Käyttämällä z_2 :ta ja liitteen 1 approksimaatiota (27) saadaan arviot

$$(4.2.37) \quad a_2 = H(fz_2)$$

$$(4.2.38) \quad b_2 = H(gz_2) \quad , \quad \text{joissa}$$

$$(4.2.39) \quad H(z) = z + \frac{1}{2} - \frac{1}{24z + \frac{2}{z+1}} \approx \psi^{-1}(\log z)$$

Laskettuja a_2 :n ja b_2 :n arvoja on liitetaulukossa 9.

Taulukosta havaitaan esimerkiksi, että arvioiden a_2 ja b_2 suhteellinen virhe on pienempi kuin 3 %, kun $a, b \geq 1$ ja pienempi kuin 0.3 %, kun $a, b \geq 2$; sekä suhteellinen että absoluuttinen virhe lähenevät nollaa, kun a ja b kasvavat.

Sen sijaan a_2 :n ja b_2 :n tarkkuus pienenee, jos a tai b on pienempi kuin 1 (jolloin $fz_1 \leq 0.5$ tai $gz_1 \leq 0.5$); näissä tapauksissa käytetään kappaleissa 6.3 tai 6.4 esitettävää menettelyä. Mainittakoon tässä, että Pearsonin taulukoista (2) löytyy vain parametrien a ja b arvoja $0.5 \leq a \leq 50$, $a \leq b \leq 50$ vastaavia betajakauman kertymäfunktioita, joten arviot ovat tarkkoja melkein kaikkien Pearsonin tarkastelemien jakaumien tapauksessa.

4.3. U-MUOTOISTEN BETAJAKAUMIEN TAPAUS: $a < 1$ JA $b < 1$

Tarkasteltavan tapauksen käsittelu perustuu likelihood yhtälöiden siirtomuotoon (2.4.31), josta seuraavat yhtälöt

$$(4.3.1) \quad \frac{1}{x} = \sqrt{\bar{u}\bar{v}}$$

$$(4.3.2) \quad \frac{1}{a} = \bar{u} + \frac{1}{x}$$

$$(4.3.3) \quad \frac{1}{b} = \bar{v} + \frac{1}{x}$$

Näissä yhtälöissä pyritään arvioimaan suuret

$$(4.3.4) \quad \bar{u} = u - u' = u - [\psi(a+b+1) - \psi(a+1)]$$

$$(4.3.5) \quad \bar{v} = v - v' = v - [\psi(a+b+1) - \psi(b+1)]$$

Korvaamalla \bar{u} ja \bar{v} vastaavasti suureilla u ja v (eli arvioimalla $u' = 0$ ja $v' = 0$) saadaan aiemmin mainittujen arvioiden (4.1.12 - 13) kanssa yhtäpitävät arviot

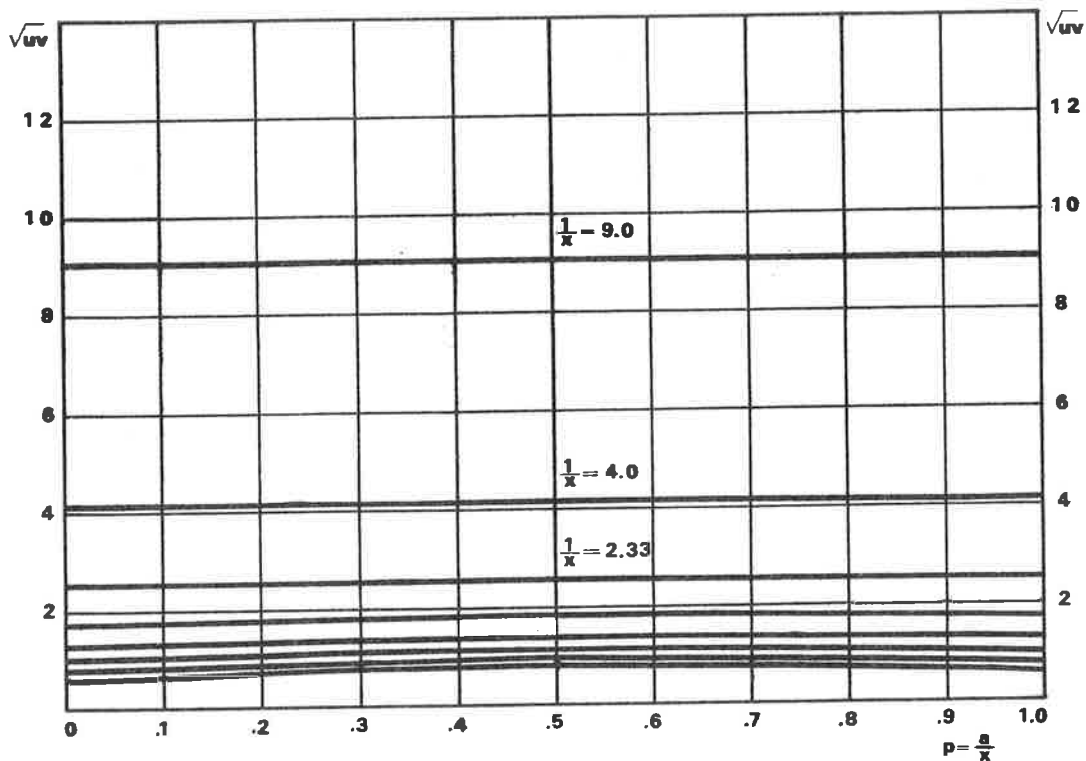
$$(4.3.6) \quad \frac{1}{x} \approx \sqrt{uv}$$

$$(4.3.7) \quad \frac{1}{a} \approx u + \sqrt{uv}$$

$$(4.3.8) \quad \frac{1}{b} \approx v + \sqrt{uv} .$$

Arvion (4.3.6) mukaan $1/x$ on pienillä $a:n$ ja $b:n$ arvoilla vain \sqrt{uv} :n funktio. Tämä riippuvuus säilyy vastaavana, vaikka a ja b kasvavat, kuten seuraavasta kuviosta havaitaan.

Diagramma V : Suureen $1/x$ nivookäyriä $p:n$ ja \sqrt{uv} :n määräämässä koordinaatistossa



Havaitaan, että $1/x$:n nivookäyrät ovat likimain horisontaalisia suoria, joten symmetrisen tilanteen $a = b$ (eli $p = 0.5$) arvioita voidaan soveltaa laajemminkin.

Kuten liitteessä 2 symmetrisen tapauksen $a = b$ avulla osoitetaan, on arviota (4.3.6) syytä korjata seuraavasti

$$(4.3.9) \quad \frac{1}{x} \approx \sqrt{uv} - \frac{1}{\sqrt{uv} + 1.5}$$

Käyttämällä identiteettiä $1/a = (1/p) \cdot (1/x)$ ja p :n arviota

(4.1.12) saadaan (4.3.9):n avulla

$$(4.3.10) \quad \frac{1}{a} \approx \left(1 + \sqrt{\frac{u}{v}}\right) \left(\sqrt{uv} - \frac{1}{\sqrt{uv} + 1.5}\right)$$

$$(4.3.11) \quad \frac{1}{b} \approx \left(1 + \sqrt{\frac{v}{u}}\right) \left(\sqrt{uv} - \frac{1}{\sqrt{uv} + 1.5}\right).$$

Jälkimmäinen arvio seuraa symmetrian nojalla.

Kun $a, b \leq 0.5$ ja $p, q \geq 0.3$ on approksimaatioiden suhteellinen virhe alle 4 %, mikä merkitsee pahimmillaan parin sadasosan absoluuttista virhettä. Kun a ja b lähestyvät nollaa häviää suhteellinen virhe nopeasti. Esimerkiksi kun $a = 0.050$ ja $b = 0.2000$ saadaan arvioiksi vastaavasti 0.0511 ja 0.2002 suhteellisten virheiden ollessa 2.2 % ja 0.1 %.

Siirrytään symmetriseen tapaukseen perustuvasta menetelmästä yleisen tapauksen tarkasteluun. Valitussa tarkastelutavassa pyritään arvioimaan suureet

$$(4.3.12) \quad u' = \psi(a+b+1) - \psi(a+1)$$

$$(4.3.13) \quad v' = \psi(a+b+1) - \psi(b+1).$$

Liitteen 1 arvio (32) voidaan kirjoittaa muotoon

$$(4.3.14) \quad \psi(x+1) \approx \log\left(1 + e^Y - \frac{0.877}{47e^Y + 2}\right).$$

Koska perusmuodon (2.4.21) perusteella on parametreille a ja b voimassa yhtälöt

$$(4.3.15) \quad e^{\psi(a)} = fe^{\psi(a+b)} = fz$$

$$(4.3.16) \quad e^{\psi(b)} = ge^{\psi(a+b)} = gz, \quad \text{jossa } z = e^{\psi(a+b)},$$

saadaan käyttämällä arviota (4.3.14) u' :lle ja v' :lle approksimaatiot

$$(4.3.17) \quad u' \approx \log\left(\frac{1 + z - \frac{0.877}{47z + 2}}{1 + fz - \frac{0.877}{47fz + 2}}\right)$$

$$(4.3.18) \quad v' \approx \log\left(\frac{1 + z - \frac{0.877}{47z + 2}}{1 + gz - \frac{0.877}{47gz + 2}}\right).$$

Kun z :n arviona käytetään jo edellisessä kappaleessa mainittua globaalista arviota z_0

$$(4.3.19) \quad z_0 = \frac{fg + d^2}{fg \cdot 2d + d^2 \left[\exp\left(\frac{uv}{\sqrt{uv + 1.2}}\right) - 1 \right]},$$

saadaan u' :lle ja v' :lle arviot u'_3 ja v'_3 :

$$(4.3.20) \quad u'_3 = \log \left| \frac{B(z_1)}{B(fz_1)} \right|$$

$$(4.3.21) \quad v'_3 = \log \left| \frac{B(z_1)}{B(gz_1)} \right|,$$

joissa $B(z) = 1 + z - \frac{0.877}{47z + 2}.$

Tämän jälkeen määrätään seuraavassa järjestyksessä suureet $\bar{u}_3, \bar{v}_3, \frac{1}{x_3}$ sekä a_3 ja b_3 :

$$(4.3.22) \quad \bar{u}_3 = u - u'_3$$

$$(4.3.23) \quad \bar{v}_3 = v - v'_3$$

$$(4.3.24) \quad \frac{1}{x_3} = \sqrt{\bar{u}_3 \bar{v}_3}$$

$$(4.3.25) \quad a_3 = \frac{1}{\bar{u}_3 + \frac{1}{x_3}}$$

$$(4.3.26) \quad b_3 = \frac{1}{\bar{v}_3 + \frac{1}{x_3}}.$$

Arviot a_3 ja b_3 ovat kaikkialla tarkkoja likiarvoja parametreille a ja b . Laskettuja arvoja on liitteessä 10.

Approksimaatioiden a_3 ja b_3 suhteellinen virhe lähestyy nollaa, kun a ja b pienenevät. Kun a ja b ovat pienempiä kuin 0.1, on a_3 :n ja b_3 :n suhteellinen virhe alle 2 %. Suhteellinen virhe

on suurimmillaan (n. 9 %), kun $x = a + b$ on likimain 0.5; tämä merkitsee pahimmassa tapauksessa a_3 :ssa ja b_3 :ssa kahden sadasosan absoluuttista virhettä. Kun $x = a + b \approx 1$, on suhteellinen virhe alle 4 % ja absoluuttinen virhe alle kaksi sadasosaa. Approksimaatioiden a_3 ja b_3 suhteellinen virhe pysyy alle 5 %:n, vaikka $x = a + b$ kasvaa¹⁾. Huomautetaan, että parametrien a ja b ollessa alle yhden, riittää tavallisesti tietty absoluuttinen tarkkuus. Käyttämällä liitteen 3 iterointiohjelmaa voidaan a ja b tarvittaessa kylläkin arvioida mielivaltaisella tarkkuudella.

Arviota a_3, b_3 huomattavasti tarkempia arvioita alueella $a \geq 2, b \geq 2$ ovat edellisessä luvussa esitetyt arviot a_1, b_1 tai a_2, b_2 . Vastaavasti alueella $a \leq 1, b \geq 3$ on edullista käyttää seuraavassa luvussa johdettuja arvioita.

1) Näin on tilanne ainakin tässä työssä taulukoiduissa pisteissä.

4.4. GAMMA-JAKAUMAN MUOTOISTEN BETA-JAKAUMIEN TAPAUS:

$$a/b \approx 0 \text{ TAI } b/a \approx 0$$

Aiemmin on jo viitattu vaikeuksiin, jotka johtuvat parametrien a ja b suuruusluokista. Koska ml-yhtälöissä

$$(4.4.1) \quad \psi(a+b) - \psi(a) = u$$

$$(4.4.2) \quad \psi(a+b) - \psi(b) = v$$

esiintyvät a , b ja $a+b$ jokainen psii-funktion argumenttina, täytyy nyt tarkasteltavassa tapauksessa käyttää eri alueiden psii-funktion arvioita. Tällöin joudutaan tavallisesti hankaliin lausekkeisiin. Seuraavassa esitetään arviointimenetelmä, jossa kuitenkin päädytään selväpiirteiseen tulokseen.

Yhtälöt (4.4.1 - 2) on syytä kirjoittaa muotoon¹⁾

$$(4.4.3) \quad \psi(b) - \psi(a) = u-v, \quad \text{jossa } \frac{a}{b} \approx 0$$

$$(4.4.4) \quad \psi(a+b) - \psi(b) = v, \quad \text{jossa } \frac{b}{a+b} \approx 1$$

ja edelleen

$$(4.4.5) \quad \log a - \psi(a) = u-v + [\log a - \psi(b)]$$

$$(4.4.6) \quad v = \psi(a+b) - \psi(b) \approx \psi'(b + \frac{a}{2}) a$$

1) Tässä tarkastelemme tapausta $a/b \approx 0$; käänteinen tapaus käsitellään vastaavasti.

Seuraavassa ei tarkemmin perustella niitä vaiheita, joiden kautta tähän menetelmään lopulta päädyttiin vaan esitetään lyhyesti tärkeimmät tulokset¹⁾.

Psii-funktion derivaatalle voidaan johtaa seuraavat asym-
toottiset kehitelmät

$$(4.4.7) \quad \psi'(x) \approx \frac{1}{x} + \frac{1}{2x^2} + \frac{1}{6x^3} \quad (\text{klassinen})$$

$$(4.4.8) \quad \psi'(x) \approx \frac{1}{x - \frac{1}{2} + \frac{1}{12x} + \frac{1}{24x^2}} .$$

Suorittamalla jako havaitaan (4.4.8):n kehitelmän vastaavan klassista arviota merkityllä tarkkuudella.

Käyttämällä (4.4.8):n 3 ensimmäistä termiä saadaan

$$(4.4.9) \quad v \approx \frac{a}{b + \frac{a}{2} - \frac{1}{2} + \frac{1}{12(b + \frac{a}{2})}} .$$

Tämän nojalla v on likipitään suoraan verrannollinen a :han ja kääntäen verrannollinen b :hen.

Samalla alueella ja vastaavalla tarkkuudella pätee

$$(4.4.10) \quad \log a - \psi(b) \approx \log\left(\frac{a}{b - \frac{1}{2} + \frac{1}{24b}}\right) .$$

1) Johto perustuu betajakauman ja gammajakauman väliseen vastaavuuteen ja sen kehittämiseen tarvittiin yli 2 vuotta!

Logaritmoitavan lausekkeen ja yhtälön (4.4.9) vasemman puolen yhdenkaltaisuus on huomionarvoinen. Kertomalla (4.4.9) auki saadaan

$$(4.4.11) \quad a \approx v \left(\frac{a}{2} \right) + v \left[b - \frac{1}{2} + \frac{1}{12 \left(b + \frac{a}{2} \right)} \right]$$

$$(4.4.12) \quad a \left[1 - \frac{v}{2} \right] \approx v \left[b - \frac{1}{2} + \frac{1}{12 \left(b + \frac{a}{2} \right)} \right]$$

$$(4.4.13) \quad \frac{a}{b - \frac{1}{2} + \frac{1}{12 \left(b + \frac{a}{2} \right)}} \approx \left(\frac{2v}{2-v} \right) .$$

Ainoa ero yhtälön (4.4.10) logaritmoitavassa ja (4.4.13):n oikeassa puolessa on termien $1/24b$ ja $1/12 \left(b + \frac{a}{2} \right)$ ero. Emme nyt kiinnitä tähän huomiota vaan arvioimme

$$(4.4.14) \quad \log a - \psi(b) \approx \log \left(\frac{2v}{2-v} \right) .$$

Kun tämä sijoitetaan yhtälöön (4.4.5), saadaan a:n määräämiseksi yhtälö

$$(4.4.15) \quad \log a - \psi(a) = M , \text{ jossa}$$

$$(4.4.16) \quad M = u-v + [\log a - \psi(b)]$$

$$\approx u-v + \log \left(\frac{2v}{2-v} \right)$$

$$= \log \left(\frac{1}{f} \right) + \log g + \log \left(\frac{2v}{2-v} \right)$$

$$= \log \left(\frac{v}{f} \right) + \log \left(\frac{2g}{2-v} \right) .$$

Gammajakauman parametrien estimoinnissa ml-menetelmällä joudutaan samaa tyyppiä olevaan yhtälöön¹⁾. Suureen M suuruusluokan, kun $a/b \approx 0$, määrää termi $\log(\frac{v}{f})$.

Arvioimalla M yhtälöstä (4.4.16), saadaan a yhtälöstä (4.4.15) ja b edelleen yhtälöstä

$$(4.4.17) \quad b \approx a \left(\frac{2-v}{2v} \right) + \frac{1}{2} - \frac{1}{6+12 \left(\frac{a}{v} \right)} .$$

Yhtälön (4.4.15) likimääräistä ratkaisua ei tässä yhteydessä johdeta, vaan tyydytään esittämään eräitä käyttökelpoisia arvioita.

Yksinkertaiset ja suhteellisen tarkat ylä- ja alarajat a:lle ovat

$$(4.4.18) \quad \frac{1}{2M} + \frac{1}{6+3M} < a < \frac{1}{2M} + \frac{1}{6+2M} .$$

Yläraja on asympotoottisesti tarkka sekä pienillä että suurilla a:n arvoilla ja sen maksimivirhe on pienempi kuin $2 \cdot 10^{-2}$.

Tarkempi arvio on

$$(4.4.19) \quad a \approx \frac{1}{2M} + \frac{1}{6+2M+\log(1+M^2)} ,$$

1) Ks. Vasama - Vartia (7) s. 509.

jonka suhteellinen virhe on suurimmillaan 1 %:n luokkaa (kun $a \approx 0,5$), jolloin myös absoluuttinen virhe saavuttaa maksimiarvonsa. Tämä on pienempi kuin $3 \cdot 10^{-3}$. Absoluuttinen virhe häviää nopeasti, jos a kasvaa tai lähenee nollaa; suhteellinen virhe jälkimmäisessä tapauksessa pysyy noin puolessa prosentissa.

Jos esimerkiksi $M = 0.5$, on $a = 1.1377$ ja arvio (4.4.19) antaa $a \approx 1.1384$, joten virhe on alle promillen.

Käytettävät approksimaatiokaavat voidaan esittää seuraavasti.

Ensin lasketaan suure

$$(4.4.20) \quad M_4 = u - v - \log\left(\frac{2-v}{2v}\right)$$

ja määrätään yhtälön

$$(4.4.21) \quad \log(a) - \psi(a) = M_4$$

likimääräinen ratkaisu a_4 muodossa

$$(4.4.22) \quad a_4 = \lambda(M_4), \quad \text{jossa}$$

$$(4.4.23) \quad \lambda(M) = \frac{1}{2M} + \frac{1}{6+2M+\log(1+M^2)}.$$

Tämän jälkeen määräytyy b :n arvio b_4 rekursiivisesti yhtälöstä

$$(4.4.24) \quad b_4 = a_4 \left(\frac{2-v}{2v} \right) + \frac{1}{2} - \frac{1}{6+12(a_4/v)} .$$

Nämä arviot toimivat erinomaisesti gammajakauman tyyppisten betajakaumien tapauksessa, jolloin a on pieni verrattuna b :hen. Kun $a \leq 1$ ja $b \geq 3$ ovat arvioiden virheet alle 1 % ja virheet häviävät nopeasti, kun a pienenee ja b kasvaa. Esimerkiksi hankalan vinossa tapauksessa $p = 0.1$, $x = 9$, jolloin $a = 0.9$ ja $b = 8.1$ saadaan arviot $a_4 = 0.89987$ ja $b_4 = 8.091$. Näiden virheet ovat vastaavasti 0.01 % ja 0.1 %.

Arviot toimivat yllättävän hyvin vaikka b olisi vain yli yhden, kun samalla $p \leq 0.3$. Erityisesti p :lle saadut arviot ovat vielä tällöin hyvin tarkkoja. Odotusarvon p virhe on pienempi kuin 1 %, mutta x :n arviolla $a_4 + b_4$ on taipumus heikentyä a :n lähetessä b :tä. Kuitenkin myös x :n virhe on pienempi kuin 2 %, jos ehdon $p \leq 0.3$ lisäksi $a \leq 1$.

Liitetaulukon 11 avulla saadaan täsmällisempi käsitys arvioiden tarkkuudesta.

Huomautetaan vielä varmuuden vuoksi, että toiseen suuntaan vinossa tapauksessa $b \leq 1$ ja $a \geq 3$ on sopivat muuttujat vaihdettava keskenään. Tällöin käytetään kaavoja

$$(4.4.25) \quad M_4 = v-u-\log\left(\frac{2-u}{2u}\right)$$

$$(4.4.26) \quad b_4 = \lambda(M_4) , \quad \text{jossa}$$

$$(4.4.27) \quad \lambda(M) = \frac{1}{2M} + \frac{1}{6+2M+\log(1+M^2)}$$

$$(4.4.28) \quad a_4 = b_4\left(\frac{2-u}{2u}\right) + \frac{1}{2} - \frac{1}{6+12(b_4/u)} .$$

Voidaan siis todeta, että saatujen kaavojen avulla pystytään arvioimaan hyvin tarkasti gammajakauman tyyppisten betajakaumien parametrit. Näin ollen luvussa 4 esitettyjen approksimaatiokaavojen avulla voidaan arvioida mm. betajakauman parametrien a ja b ja geometristen keskiarvojen f ja g väliset yhteydet riippumatta $a:n$ ja $b:n$ suuruusluokista.

KIRJALLISUUSVIITTEET

M. Abramowitz & I.A. Stegun (9): Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables, Dover Publications, Inc. New York.

F.J. Anscombe (5): Testing to Establish a High Degree of Safety or Reliability, Bulletin of Statistical Institute, 39, 73-89, 1961.

B. Baldessari (4): Estimate of the Parameters in the Gini's Hypothesis, Statistical Theory and Method Abstract, 10/796, 1966.

H. Bateman (17): Higher Transcendental Functions, McGraw - Hill Company, Inc. 1953.

H. Cramér (11): Mathematical Methods of Statistics, Princeton University Press, USA 1946.

W.P. Elderton and N.L. Johnson (3): Systems of Frequency Curves, Cambridge University Press, 1969.

W.Fulks (10): Advanced Calculus, John Wiley & Sons, Inc. New York. London, Sydney, Sixth printing 1965.

R. Gnadesikan, R.S. Pinkham and Laura P. Hughes (Bell Telephone Laboratories, Inc.) (1): Maximum Likelihood Estimation of the Parameters of the Beta Distribution from Smallest Order Statistics, Technometrics Vol. 9, No. 4, 607-620, November 1967.

P. Henrici (19): Elements of Numerical Analysis, John Wiley & Sons, Inc. New York. London 1964.

K.V. Laurikainen (16): Modernin fysiikan alkeet, WSOY, Porvoo. Helsinki 1968.

Rolf Nevanlinna (15): Suhteellisuusteorian periaatteet, WSOY, Porvoo 1964.

K. Pearson (2): Tables of the Incomplete Beta-function, Cambridge University Press, 1968 (First Edition 1934).

C.R. Rao (6): Linear Statistical Inference and its Applications, John Wiley & Sons, Inc. New York. London. Sydney 1965.

W. Rudin (14): Principles of Mathematical Analysis, Kogakusha Company Ltd. Tokyo 1964.

L. Törnqvist, L. Nordberg (13): Päättätieteen keskeisiä ongelmia, WSOY, Porvoo. Helsinki 1968.

P.L.I. Vartia & Y.O. Vartia (8): F-distribution as a Model for Income Distribution, Institute of Economics, University of Helsinki, Research Reports No. 21, Helsinki 1972.

Y.O. Vartia (18): Betajakauman parametrien maximum likelihood estimaattoreiden approksimoinnista. Tilastotieteen pro gradu työ, Valtiotieteellinen tiedekunta, Helsinki 1971.

P.M. Vasama, Y. Vartia (7): Johdatus tilastotieteeseen, Ylioppilastuki ry, Helsinki 1970.

Samuel S. Wilks (12): Mathematical Statistics, John Wiley & Sons, Inc. New York, London 1961.

LIITE 1: PSII-FUNKTION JA PSII-FUNKTION KÄÄNTEISFUNKTION
APPROKSIMOINTI

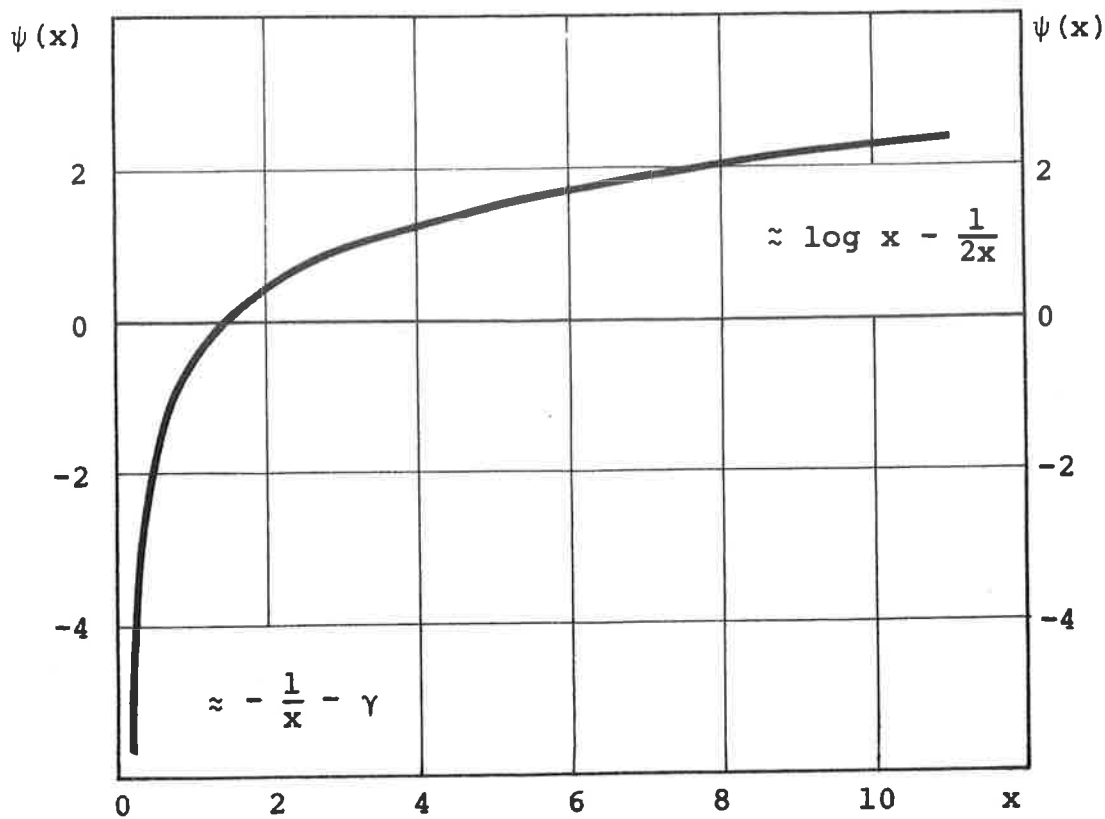
Maximum likelihood-yhtälöiden ratkaisemiseksi täytyy hallita psii-funktion käyttäytyminen, jonka vuoksi tässä johdetaan erityisesti kyseiseen ongelmaan soveltuvia uusia psii-funktion approksimointikaavoja. Psii-funktion käänteisfunktioon perustuva ml-yhtälöiden kirjoitustapa osoittautuu myös hyvin käyttökelpoiseksitehtävää ratkaistaessa. Tällöin joudutaan johtamaan ilmeisesti uusia psii-funktion käänteisfunktiota koskevia tuloksia. Psii-funktion käänteisfunktiota ei tietääkseni ole aikaisemmin tarkasteltu itsenäisenä funktiona.

Maximum likelihood-yhtälöiden ratkaisemisessa joudutaan oleellisesti erilaisiin ratkaisuihin suurilla ja pienillä $a:n$ ja $b:n$ arvoilla. Tämä johtuu pohjimmiltaan psii-funktion eri tyyppisestä käyttäytymisestä pienillä ja suurilla argumentinsa arvoilla. Nimittäin $\psi(x) \approx \log x$, kun x on suuri, kun taas pienillä positiivisilla $x:n$ arvoilla $\psi(x) \approx -\left(\frac{1}{x}\right)$. Näiden tietojen samanaikainen huomioonottaminen on yllättävän hankalaa; voi jopa sanoa, että tämä on tutkimuksen perusongelma. Psii-funktio määritellään seuraavasti:

$$(1) \quad \psi(x) = \frac{d}{dx} \log \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}, \text{ jossa}$$

$$(2) \quad \Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt = \text{Eulerin gammafunktio.}$$

Diagramma 1. Psii-funktio positiivisilla argumentin arvoilla



Taulukko 1. Psii-funktion arvoja

x	$y = \psi(x)$
0.01	-100.5609
0.1	- 10.4238
0.2	- 5.2890
0.3	- 3.5025
0.4	- 2.5614
0.5	- 1.963510
1.0	- 0.577216
1.5	+ 0.036490
2.0	+ 0.422784
3.0	0.922784
4.0	1.256118
5.0	1.506118
6.0	1.706118

$$\psi(1) = -\gamma$$

$$\psi^{-1}(0) = 1.4616$$

Erityisesti on syytä mainita psii-funktion tärkeä siirto-ominaisuus

$$(3) \quad \psi(x+1) = \psi(x) + \frac{1}{x} \quad ,$$

joka seuraa gammafunktion kertomaominaisuudesta

$$(4) \quad \Gamma(x+1) = \Gamma(x) \cdot x \quad .$$

Psii-funktiosta on syytä todeta seuraavat klassiset tulokset:

$$(5) \quad \begin{aligned} \psi(x) &= \log x - \frac{1}{2x} - \frac{1}{12x^2} + \frac{1}{120x^4} + o\left(\frac{1}{x^6}\right) \\ &= \log x - \frac{1}{2x} - \sum_{n=1}^m \frac{B_{2n}}{2n} \left(\frac{1}{x}\right)^{2n} + o\left(\frac{1}{x^{2m+2}}\right) \quad , \end{aligned}$$

jossa B_{2n} :t ovat Bernoullin lukuja.

$$(6) \quad \psi(x+1) = -\gamma - \sum_{n=1}^{\infty} \zeta(n+1) (-x)^n \quad , \quad \text{kun } |x| < 1 \quad , \quad \text{jossa}$$

$$(7) \quad \gamma = \text{Eulerin vakio} = -\psi(1) \approx 0.577216 \quad \text{ja}$$

$$(8) \quad \zeta(z) = \sum_{k=1}^{\infty} \frac{1}{k^z} = \text{Riemannin zetafunktio}$$

$$(9) \quad \psi(x) = -\gamma - \left(\frac{1}{x}\right) + \sum_{n=1}^{\infty} \frac{x}{n(x+n)} \quad .$$

Bateman (17) esittää näiden lisäksi useita muitakin mielenkiintoisia tuloksia. Hyvä taulukko on löydettävissä teoksesta Abramowitz & Stegun (9).

Seuraavat uudet approksimointikaavat on johdettu kaavan (5) avulla. Siirtämällä termejä logaritmin sisään saadaan¹⁾ suurilla x :illä käyttökelpoiset kaavat:

$$(10) \quad \psi(x) = \log\left(x - \frac{1}{2} + \frac{1}{24x} + \frac{1}{48x^2} + \frac{23}{5740x^3} + o(x^{-4})\right)$$

$$(11) \quad \psi(x) = \log\left(x - \frac{1}{2} + \frac{1}{24\left(x - \frac{1}{2}\right)} + o(x^{-3})\right)$$

Kehittämällä kaavassa (5) $\frac{1}{x}$:n potenssisarja ketjumurtoluvuksi saadaan esimerkiksi tietokonekäyttöön hyvin soveltuva uusi kaava:

$$(12) \quad \psi_2(x) = \log x - \frac{A(x)}{2x},$$

$$\text{jossa} \quad A(x) = 1 + \frac{1}{6x + \frac{6}{10x + \left(\frac{6}{1 + 1.2879x}\right)}}.$$

1) Yliopistollisissa ja tieteellisissä julkaisuissa on tapana käyttää persoonattomia ilmaisuja (esim. passiivia) myös sellaisissa tilanteissa, joissa kirjoittajan henkilökohtainen panos on ilmeinen. Esimerkiksi eo. ilmaisu 'saadaan suurilla x :illä käyttökelpoiset kaavat' tulisi paremminkin olla 'minä olen pohdittuani eri vaihtoehtoja päättänyt seuraaviin suurilla x :illä käyttökelpoisiin kaavoihin'. Tulen jatkossakin noudattamaan vallitsevaa vaikkakin epätasällistä käytäntöä.

Kaava on sovitettu kun $x=1$ ja se antaa oikean asymptoottisen kehitelmän termiin $\frac{1}{120x^4}$ saakka. Lisäksi kaava käyttäytyy likimain kuin $-\left(\frac{1}{x}\right)$, kun $x \rightarrow 0$. Tarkkuus on suurempi kuin 6 desimaalia, kun $x \geq 1$ ja vähintään 2 merkitsevää numeroa, kun $0 < x < 1$. Käyttämällä siirto-ominaisuutta $\psi(x) = \psi(x+1) - \frac{1}{x}$ saadaan erinomainen kaava myös alueella $0 < x \leq 1$. Joitakin laskettuja arvoja on sivulla 94.

Arvio (12) on parempi kuin sovellutuksissa aikaisemmin käytetyt arviot. Esimerkiksi artikkelissa (1) käytetyssä $\psi(x)$:n approksimaatiokaavassa jouduttiin vastaavan tarkkuuden saavuttamiseksi käyttämään siirto-ominaisuutta (5.2) 12 kertaa, mikä itse asiassa merkitsee funktion $\psi(x+13)$ approksimointia, kun $x \geq 0$.

Edellisiä epätarkempi, mutta hyvin yksinkertainen globaalinen psii-funktion approksimaatiokaava on

$$(13) \quad \psi_3(x) = \log x - \frac{1}{x} \left(\frac{1 + 3x}{1 + 6x} \right) .$$

Tämä on kahden merkitsevän numeron tarkkuudella oikea, kun $0 < x \leq 2$ ja tarkentuvasti oikea, kun $x \rightarrow \infty$. Laskettuja arvoja on sivulla 94.

Psii-funktion globaalinen approksimointikaava

$$(14) \quad \psi_4(x) = \log \left(\frac{1}{e^{\frac{1}{x}} - 1} \right)$$

on kaikessa kummallisuudessaan hyvin paljastava: se sisältää useita erilaisia vääntöjä ja vastavääntöjä, joita kuvaavat erilaiset funktiot ja niiden käänteisfunktiot kombinoituina eri järjestyksiin.

Määritellään seuraavat merkinnät:

$$(15) \quad \text{inv}(t) = \frac{1}{t} = \text{inv}^{-1}(t)$$

$$(16) \quad \text{exp}(t) = e^t = \log^{-1}(t) .$$

Tällöin voidaan (1) kirjoittaa operatiiviseen muotoon

$$(17) \quad \psi_4(x) = \log\{\text{inv}(\text{exp}[\text{inv}(x)] - 1)\},$$

josta näkyy eri vääntöjen mutkikas yhteispeli.

Tämä kaava johdetaan yksikertaisesti käänteisfunktiokaavan (33) avulla. Sivulla 94 olevasta taulukosta havaitaan, että $\psi_4(x)$ on kaikilla x :n arvoilla lievä yliampuja.

Eräiden psii-funktion approksimaatiokaavojen tarkkuudet

x	$\psi(x)$	$\psi_2(x)$	$\psi_3(x)$	$\psi_4(x)$
0.1	-10.4238	-10.5 -	-10.428 -	-10.0 +
0.2	- 5.28904	- 5.31 -	- 5.25 +	- 5.0 +
0.3	- 3.50253	- 3.507 -	- 3.47 +	- 3.3 +
0.4	- 2.56139	- 2.563 -	- 2.53 +	- 2.4 +
0.5	- 1.96351	- 1.9639 -	- 1.94 +	- 1.85 +
0.6	- 1.54062	- 1.5407 -	- 1.53 +	- 1.46 +
0.7	- 1.22003	- 1.2206 -	- 1.21 +	- 1.15 +
0.8	- 0.965017	- 1.965017	- 0.956 +	- 0.91 +
0.9	- 0.754929	- 0.754929	- 0.748 +	- 0.71 +
1.0	- 0.577216	- 0.577216	- 0.571 +	- 0.54 +
1.1	- 0.423755		- 0.419 +	- 0.39 +
1.2	- 0.289041	tarkkuus	- 0.285 +	- 0.26 +
1.3	- 0.169192	yli 6	- 0.166 +	- 0.14 +
1.4	- 0.061386	desimaalia	- 0.059 +	- 0.04 +
1.5	+ 0.036489		+ 0.039 +	+ 0.05 +
2.0	+ 0.422784		+ 0.424 +	+ 0.43 +
2.5	0.703156		0.7038 +	
3.0	0.922784		0.9232 +	0.927 +
3.5	1.10316		1.1034 +	
4.0	1.25612		1.2563 +	1.259 +
4.5	1.38887		1.3890 +	
5.0	1.50612		1.5062 +	1.508 +
5.5	1.61109		1.6112 +	
6.0	1.70612		1.70617 +	1.707 +

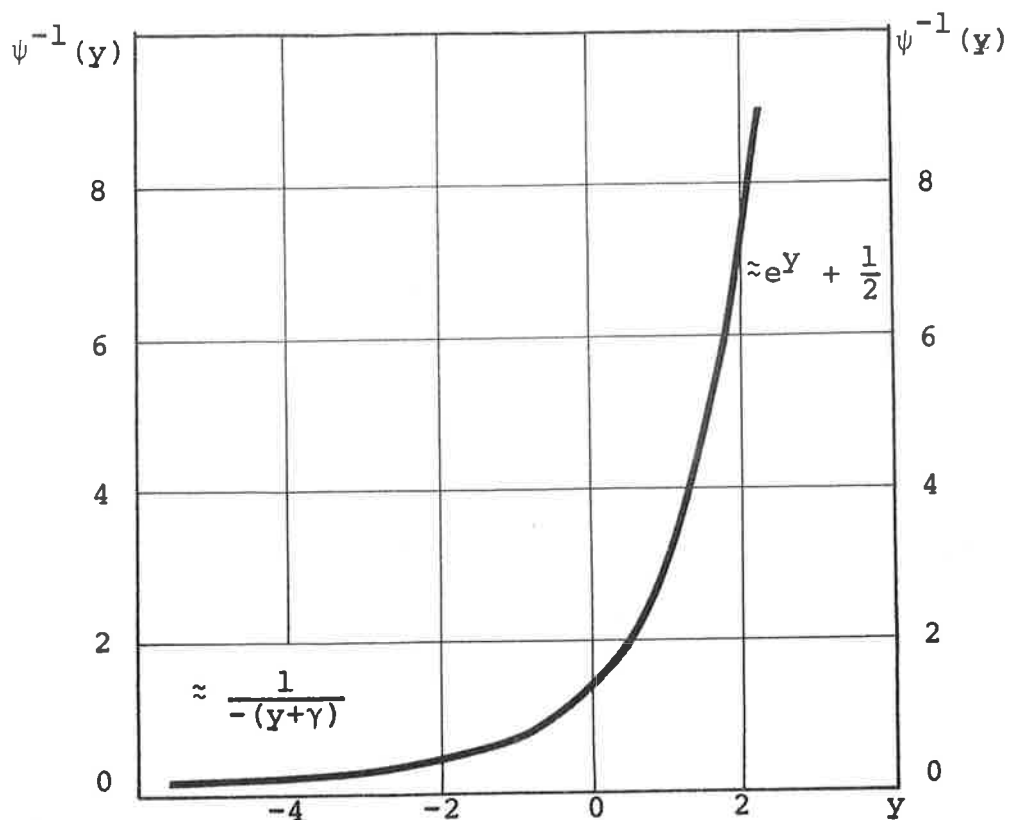
$$\psi_2(x) = \log x - \frac{1}{2x} \left\{ 1 + \frac{1}{6x + \frac{6}{10x + \left(\frac{6}{1 + 1.2879x}\right)}} \right\}$$

$$\psi_3(x) = \log x - \frac{1}{x} \left(\frac{1+3x}{1+6x} \right)$$

$$\psi_4(x) = \log \left(\frac{1}{\frac{1}{e^x} - 1} \right)$$

Sarakkeissa on tarkkuuden havainnollistamiseksi merkitty oikeat desimaalit sekä 1. väärä desimaali pyöristettynä. Lukuarvojen jälkeen on merkitty, onko approksimaatio yliampuja (+) vai aliampuja (-) kyseisessä pisteessä.

Diagramma 2. Psii-funktion käänteisfunktio



Taulukko 2. Psii-funktion käänteisfunktion arvoja

y	$\psi^{-1}(y)$
-100	0.0100563
- 50	0.0202202
- 20	0.051271
- 10	0.104357
- 8	0.131231
- 5	0.211614
- 4	0.26392
- 3	0.34689
- 2	0.49270
- 1	0.78500
0	1.46163
1	3.2032
2	7.88343
3	20.58347
4	55.09739
5	148.91288

Koska psii-funktio on kaikkialla monotonisesti kasvava, on sillä myös monotonisesti kasvava käänteisfunktio. Merkitään

$$(18) \quad y = \psi(x) \text{ , jolloin siis}$$

$$(19) \quad x = \psi^{-1}(y) .$$

Muuttujan x vaihtelualue on positiivinen reaaliakseli, kun taas y voi saada mielivaltaisen reaalisen arvon. Psii-funktion käänteisfunktion approksimointia negatiivisilla y :n arvoilla ei tässä yhteydessä ole syytä tarkastella¹⁾, joten siirrytään psiiin käänteisfunktion tarkasteluun, kun y on positiivinen ja x siis melkoisen suuri.

Korotetaan (11) puolittain e :n potenssiin:

$$(20) \quad e^y \approx x - \frac{1}{2} + \frac{1}{24(x - \frac{1}{2})} .$$

Tästä saadaan sijoittamalla $e^y \approx x - \frac{1}{2}$ ja sieventämällä

$$(21) \quad x = \psi^{-1}(y) \approx e^y + \frac{1}{2} - \frac{1}{24e^y} .$$

1) Perustulokset tältä alueelta löytyvät tilastotieteen pro-gradu tutkimuksestani "Betajakauman parametrien maximum-likelihood estimaattoreiden approksimoinnista", Helsingin yliopisto, 1971.

Tämän arvion tarkkuudesta saa kuvan sivun 101 taulukosta. Vastaavaa menetelmää edelleen soveltamalla saadaan johdetuksi lisää termejä psiin käänteisfunktion Laurent-kehityksen e^y :ssä:

$$(22) \quad \psi^{-1}(y) = e^y + \frac{1}{2} - \frac{1}{24} e^{-y} + 0 \cdot e^{-2y} + \lambda e^{-3y} + o(e^{-4y}),$$

jossa $\lambda = 0.00467359$.

Nämä kaavat käyttäytyvät virheellisesti negatiivisilla y :n arvoilla. Seuraavissa arvioissa on näitä virheitä korjattu muuttamatta oleellisesti käyttäytymistä positiivisilla y :n arvoilla. Jos käytetään kaavan (22) kolmea ensimmäistä termiä ja pakotetaan arvio ehtoon $\psi^{-1}(-\infty)=0$, saadaan erääksi mahdolliseksi arvioksi

$$(23) \quad \psi^{-1}(y) \approx e^y + \frac{1}{2} - \frac{1}{24e^y + 2}.$$

Tämän tarkkuus on yli 3 desimaalia, kun $y \geq 0$. Toinen mahdollisuus on yhdistää kaavan (22) toinen ja kolmas termi yhdeksi geometrisen sarjan avulla, jolloin saadaan

$$(24) \quad \psi^{-1}(y) \approx e^y + \frac{6e^y}{12e^y + 1} = e^y \left(1 + \frac{6}{12e^y + 1}\right).$$

Arvion (24) kehityksen e^y :ssä yhtyy (22):een termiin $-\frac{1}{24} e^{-y}$ saakka ja sen tarkkuus on kaikkialla samaa suuruusluokkaa kuin kaavan (23).

Käyttämällä hyväksi approksimoinnissa usein hyvin käyttökelpoisia ketjumurtokehitelmiä, saadaan kaavaa (22) vastaavia kehitelmiä. Seuraava kehitelmä antaa lisäinformaatiota arvion (23) asymptoottisista ominaisuuksista:

$$(25) \quad \psi^{-1}(y) = e^y + \frac{1}{2} - \frac{1}{24e^y + \frac{\lambda_1}{e^y}} + o(e^{-4y}),$$

jossa $\lambda_1 = 576\lambda = 2.691988$.

Kaavan päätermin arvoja on sivulla 101.

Vastaavasti seuraava kaava kertoo arvion (24) asymptoottisesta käyttäytymisestä:

$$(26) \quad \psi^{-1}(y) = e^y + \frac{1}{2 + \frac{1}{6e^y + \frac{\lambda_2}{e^y}}} + o(e^{-4y})$$

$$= e^y + \frac{6e^y}{12e^y + \frac{1}{1 + \frac{\lambda_2}{6e^{2y}}}} + o(e^{-4y}),$$

jossa $\lambda_2 = 144\lambda = 0.672997$.

Kaavasta (25) saadaan jatkon kannalta erittäin käyttökelpoinen globaalinen kaava korjaamalla sitä seuraavasti:

$$(27) \quad \psi^{-1}(y) \approx x_5 = e^y + \frac{1}{2} - \frac{1}{24e^y + \frac{2}{e^y + 1}} .$$

Tässä arviossa yhdistyvät arvioiden (23) ja (25) hyvät puolet; sen arvoja on laskettu sivulla 101. Saatu kaava on erinomainen, kun $y \geq 1$. Tällöin se antaa vähintään 3 oikeata desimaalia ja virhe häviää nopeasti y :n kasvaessa.

Globaalinen psiin käänteisfunktion kaava voidaan kehittää seuraavasti. Psi-funktion approksimointikaavan (10) perusteella suurilla x :n arvoilla

$$(28) \quad e^{\psi(x)} = e^y \approx x - \frac{1}{2} + \frac{1}{24x} + \frac{1}{48x^2} .$$

Saman kaavan (10) ja geometrisen sarjan summakaavan perusteella saadaan

$$(29) \quad \psi(x+1) \approx \log\left(x + \frac{1}{2} + \frac{1}{24(x+1)} + \frac{1}{48(x+1)^2}\right) \\ \approx \log\left[1 + \left(x - \frac{1}{2} + \frac{1}{24x}\right) + \frac{1}{48x^2} - \alpha\left(\frac{1}{x^2}\right)\right] ,$$

jossa $\alpha = \left(\frac{1}{24}\right)^2$.

Kun tähän sijoitetaan (28) ja näin saatu arvio sijoitetaan identiteettiin (30)

$$(30) \quad x = \frac{1}{-y + \psi(1+x)} , \text{ saadaan}$$

$$(31) \quad x \approx \frac{1}{-y + \log(1 + e^y - \frac{1}{(24)^2 e^{2y}})}$$

Tämän arvion virhe on pieni vain, kun y on positiivinen. Korjaamalla arvion käyttäytymistä, kun $y \rightarrow -\infty$ ja pakottamalla se pisteiden $y = 0$ ja $y = -2$ lähistölle, päädyin seuraavaan melkoisen tarkkaan ja yksinkertaiseen globaaliseen kaavaan:

$$(32) \quad x \approx \frac{1}{-y + \log(1 + e^y - \frac{0.877}{47e^y + 2})}$$

Saatu lauseke käyttäytyy samoin kuin $\frac{1}{-(y+\gamma)}$, kun $y \rightarrow -\infty$ ja sen maksimivirhe on noin $1 \cdot 10^{-2}$, jonka se saa kun $x \approx 0.2$ tai $x \approx 3.2$.

Yksinkertaistamalla kaavaa (32) saadaan mielenkiintoinen globaalinen kaava (33), jonka virhe on aina pienempi kuin $3 \cdot 10^{-2}$:

$$(33) \quad x \approx \frac{1}{-y + \log(1 + e^y)}$$

$$= \frac{1}{\log(1 + \frac{1}{e^y})}$$

Saatu lauseke on aina lievä aliampuja. Arvion avulla saadaan johdettua psii-funktion approksimaatiokaava (14) ratkaisemalla se muuttujan $y = \psi(x)$ suhteen, jolloin saadaan vastavasti psii-funktiolle yliampuja.

Eräiden psii-funktion käänteisfunktion approksimaatiokaavojen tarkkuudet

y	$x = \psi^{-1}(y)$	x_4	x_5	x_6
-100	0.0100563			
- 50	0.0202202			
- 20	0.051271			
- 10	0.104357			
- 8	0.131231		0. -	
- 5	0.211614	- 6. -	0. -	
- 4	0.26392	- 2.0 -	0.1 -	
- 3	0.34689	- 0.3 -	0.2 -	
- 2	0.49270	+ 0.3 -	0.44 -	
- 1	0.78500	+ 0.75 -	0.77 -	
0	1.46163	1.458 -	1.460 -	1.463 +
1	3.2032	3.2030 -	3.2031 -	3.2032 +
2	7.88343	7.88342 -	7.88343 -	7.883429
3	20.58347	20.58343 -	20.58347 -	20.583466
4	55.09739	55.09739 -	55.09739 -	55.097387
5	148.91288	148.91288 -	148.91288 -	148.912884

$$x_4 = e^y + \frac{1}{2} - \frac{1}{24e^y}$$

$$x_5 = e^y + \frac{1}{2} - \frac{1}{24 e^y + \frac{2}{e^y + 1}}$$

$$x_6 = e^y + \frac{1}{2} - \frac{1}{24 e^y + \frac{2.601988}{e^y}}$$

Sarakkeissa on tarkkuuden havainnollistamiseksi merkitty oikeat desimaalit sekä 1 väärä desimaali pyöristettynä. Lukuarvojen jälkeen on merkitty, onko approksimaatio yliampuja (+) vai aliampuja (-) kyseisessä pisteessä.

LIITE 2. ML-YHTÄLÖIDEN RATKAISU SYMMETRISISSÄ TAPAUKSESSA $a = b$

Tarkastellaan ensin tapausta $x=a+b$ on pieni, jolloin on syytä lähteä yhtälöstä

$$(1) \quad \psi(x) - \psi\left(\frac{x}{2}\right) = u = \sqrt{uv} .$$

Tämä voidaan kirjoittaa yhtälöä (2.4.31) vastaavaan muotoon

$$(2) \quad \frac{b}{a(a+b)} = \frac{1}{x} = \sqrt{uv} - u' , \text{ jossa}$$

$$(3) \quad u' = \psi(x+1) - \psi\left(\frac{x}{2}+1\right) \\ \approx \log\left(x+\frac{1}{2}\right) - \log\left(\frac{x}{2}+\frac{1}{2}\right) .$$

Arvio ei ole erityisen tarkka, mutta on käyttökelpoinen.

Soveltamalla logaritmien erotukseen väliarvolauseesta saatavaa arviota $f(x+h) - f(x) \approx f'(x+\frac{h}{2}) \cdot h$ saadaan

$$(4) \quad u' \approx \frac{\frac{1}{2}x}{\frac{1}{2} + \frac{3}{4}x} = \frac{x}{1 + \frac{3}{2}x}$$

$$(5) \quad \frac{1}{x} \approx \sqrt{uv} - \frac{x}{1 + \frac{3}{2}x}$$

$$(6) \quad \frac{1}{x} \approx \sqrt{uv} - \frac{1/\sqrt{uv}}{1 + \frac{3}{2}/\sqrt{uv}}$$

$$(7) \quad \frac{1}{x} \approx \sqrt{uv} - \frac{1}{\sqrt{uv} + 1.5} .$$

Arviossa (6) on x :ää arvioitu $1/\sqrt{uv}$:llä. Arvion (7) tarkkuus ilmenee seuraavasta taulukosta

(8)

$\frac{1}{x}$	tapaus $p = \frac{1}{2}$	tapaus $p = 0.1$
9	8.988	8.986
4	3.981	3.972
2.333	2.326	2.304
1.5	1.513	1.468
1.0	1.040	0.969

Havaitaan, että (7) on huomattava parannus karkealle arviolle $\frac{1}{x} \approx \sqrt{uv}$. Kombinoimalla (7) yksinkertaisiin p :n ja q :n arvioihin saadaan kaavat

$$(9) \quad \frac{1}{a} \approx (1 + \sqrt{\frac{u}{v}}) \left(\sqrt{uv} - \frac{1}{\sqrt{uv} + 1.5} \right)$$

$$(10) \quad \frac{1}{b} \approx (1 + \sqrt{\frac{v}{u}}) \left(\sqrt{uv} - \frac{1}{\sqrt{uv} + 1.5} \right)$$

Seuraavaksi tarkastellaan tapausta $x = a+b$ on suuri, jolloin on syytä lähteä yhtälöstä

$$(11) \quad e^{\psi(\frac{x}{2})} / e^{\psi(x)} = f = \frac{f+g}{2} .$$

Käytetään arviota

$$(12) \quad e^{\psi(x)} \approx x - \frac{1}{2} + \frac{1}{24x}$$

$$(13) \quad \frac{x}{2} - \frac{1}{2} + \frac{1}{24\left(\frac{x}{2}\right)} \approx \left(\frac{f+g}{2}\right) \left(x - \frac{1}{2} + \frac{1}{24x}\right),$$

$$(14) \quad \frac{x}{2}(1-f-g) \approx \frac{1}{2} - \frac{1}{12x} + \left(\frac{f+g}{2}\right) \left(-\frac{1}{2} + \frac{1}{24x}\right)$$

Käyttämällä merkintää $d = 1-f-g$ saadaan

$$(15) \quad xd \approx 1 - \frac{1}{6x} + (1-d) \left(-\frac{1}{2} + \frac{1}{24x}\right)$$

$$\approx \frac{1}{2} - \frac{3}{24x} + d \left(\frac{1}{2} - \frac{1}{24x}\right).$$

Sijoittamalla vasemmalle puolelle $x \approx \frac{1}{2d}$ ja jakamalla d :llä saadaan

$$(16) \quad x \approx \frac{1}{2d} - \frac{6}{24} + \frac{1}{2} - \frac{2d}{24}$$

$$= \frac{1}{2d} + \frac{1}{4} - \frac{d}{12}.$$

Vastaavien, mutta mutkikkaampien laskelmien avulla, jossa

(12):n tilalla käytetään tarkempaa arviota

$$(17) \quad e^{\psi(x)} \approx x - \frac{1}{2} + \frac{1}{24x} + \frac{1}{48x^2},$$

voidaan osoittaa, että (16):ssa oleva d :n kerroin on väärä.

Asymptoottisesti oikea arvio on

$$(18) \quad x \approx \frac{1}{2d} + \frac{1}{4} - \frac{d}{2} ,$$

jonka virhe symmetrisessä tapauksessa on pienempi kuin 10^{-2} , kun $x \geq 4$ tai $d \leq 6 \cdot 10^{-2}$. Tätäkin käyttökelpoisempi arvio on

$$(19) \quad x \approx \frac{1-d^2}{2d-d^2} ,$$

jonka asymptoottinen kehitelmä

$$(20) \quad \frac{1-d^2}{2d-d^2} \approx \frac{1}{2d} + \frac{1}{4} - \frac{3d}{8}$$

ei tosin ole täysin 'oikea'. Arvio (19) on symmetrisessä tapauksessa tarkka, kun $x = 1$ ja sen virhe on pienempi kuin $3 \cdot 10^{-2}$, kun $x \geq 1$. Kombinoimalla (19) yksinkertaisiin p:n ja q:n arvioihin saadaan kaavat

$$(21) \quad a \approx \left(\frac{f+d}{1+d}\right) \left(\frac{1-d^2}{2d-d^2}\right) = \frac{(f+d)(1-d)}{d(2-d)}$$

$$b \approx \left(\frac{g+d}{1+d}\right) \left(\frac{1-d^2}{2d-d^2}\right) = \frac{(g+d)(1-d)}{d(2-d)} .$$

Symmetrisen tapauksen ratkaisusta saadaan mielenkiintoista lisävalaistusta tarkastelemalla likelihödfunktiota ehdolla $a = b$. Likelihödfunktion logaritmillemme $l(a,b)$ pätee

$$(23) \quad \frac{1}{n} l(a, b) = \log \Gamma(a+b) - \log \Gamma(a) - \log \Gamma(b) \\ (a-1) \frac{1}{n} \sum \log x_i + (b-1) \frac{1}{n} \sum \log(1-x_i),$$

josta spesifioimalla $a = b$ saadaan

$$(24) \quad \frac{1}{n} l(a, a) = \log \Gamma(2a) - 2 \log \Gamma(a) - (a-1)(u+v).$$

Derivoimalla $a:n$ suhteen

$$(25) \quad \frac{d}{da} \frac{1}{n} l(a, a) = 2\psi(2a) - 2\psi(a) - (u+v)$$

ja merkitsemällä $x = 2a$ saadaan symmetrisen tapauksen
likelihood-yhtälö

$$(26) \quad \psi(x) - \psi\left(\frac{x}{2}\right) = \left(\frac{u+v}{2}\right).$$

Havaitaan, että (26) saadaan myös yleisen tapauksen likelihood-
yhtälöistä

$$(27) \quad \psi(x) - \psi(a) = u$$

$$(28) \quad \psi(x) - \psi(b) = v$$

yhteenlaskemalla ja spesifioimalla $a = b$. Toinen kirjoitus-
muoto symmetrisen tapauksen likelihood-yhtälölle on

$$(29) \quad e^{\psi\left(\frac{x}{2}\right)} / e^{\psi(x)} = \sqrt{fg}.$$

Havaitaan, että symmetriseen tapaukseen sovellettu (sidottu) maximum likelihood-menetelmä ehdottaa, että tapauksessa $a = b$ tulisi $x = a+b$ lausua seuraavien suureiden funktiona:

$$(30) \quad \left(\frac{u+v}{2}\right) = u:n \text{ ja } v:n \text{ aritmeettinen keskiarvo} \\ \text{(pienet } x:t)$$

$$(31) \quad \sqrt{fg} = f:n \text{ ja } g:n \text{ geometrinen keskiarvo} \\ \text{(suuret } x:t)$$

Voimmekin todeta, että sidottu maximum likelihood-menetelmä johtaa harhaan, sillä $\left(\frac{u+v}{2}\right)$ tai \sqrt{fg} eivät ole edes likimain vakioita kiinteillä x :n arvoilla mikäli ehto $a = b$ ei ole voimassa. Tämä osoittaa, että sidotussa tapauksessa $a = b$ hyvät kaavat (26) ja (29) eivät ole yleisesti käyttökelpoisia. Tulos on osoitus siitä, että maximum likelihood-menetelmä luottaa tosikkomaisesti tehtyihin oletuksiin ja voi antaa aivan virheellisiä ja omituisia tuloksia, mikäli oletukset ovat vääriä.

Yleisen tapauksen tarkastelujen nojalla tiedetään, että (30):n ja (31):n aritmeettiset ja geometriset keskiarvot pitää kääntää: pienillä x :n arvoilla tulee suorittaa korvaus

$$(32) \quad \left(\frac{u+v}{2}\right) \rightarrow \sqrt{uv}$$

ja suurilla x :n arvoilla korvaus

$$(33) \quad \sqrt{fg} \rightarrow \left(\frac{f+g}{2}\right) = \left(\frac{1-d}{2}\right) .$$

Symmetrisen tilanteen 'maximum likelihood-yhtälöksi' tulee valita (spesifioimalla yleisen tapauksen ml-yhtälöistä) yhtälöiden (26) ja (29) sijaan yhtälöt

$$(34) \quad \psi(x) - \psi\left(\frac{x}{2}\right) = \sqrt{uv} , \text{ kun } x \text{ on pieni}$$

$$(35) \quad e^{\psi\left(\frac{x}{2}\right)} / e^{\psi(x)} = \left(\frac{f+g}{2}\right) , \text{ kun } x \text{ on suuri.}$$

Juuri näiden yhtälöiden avulla on edellä johdettu erityisesti symmetrisessä tapauksessa tarkat kaavat.

Yhteenveto: Ongelmana on valita aritmeettisen ja geometrisen keskiarvon väliltä. Päätöstä ei saa perustaa sidottuun ml-yhtälöön (jota ei tässä pyritä ratkaisemaan) vaan ongelman laajempaan tarkasteluun.

LIITE 3. ML-YHTÄLÖIDEN ITERATIIVINEN RATKAISU JA SIIHEN
LIITTYVÄ TIETOKONEOHJELMA

Käyttämällä Newtonin iterointimenetelmää saadaan likelihood-
yhtälöille

$$(1) \quad \psi(a+b) - \psi(a) = u$$

$$(2) \quad \psi(a+b) - \psi(b) = v$$

numeerinen ratkaisu jokaiselle parille $u > 0$ ja $v > 0$.

Käytämme samoja merkintöjä kuin Henrici (19) s. 105.

Ratkaistavat yhtälöt ovat

$$(3) \quad F(a,b) = \psi(a+b) - \psi(a) - u = 0$$

$$(4) \quad G(a,b) = \psi(a+b) - \psi(b) - v = 0 .$$

Arvioimme F :ää ja G :tä Taylor-polynomilla pisteessä (a,b) :

$$(5) \quad F(a+\delta, b+\epsilon) \approx F(a,b) + F_a(a,b)\delta + F_b(a,b)\epsilon$$

$$(6) \quad G(a+\delta, b+\epsilon) \approx G(a,b) + G_a(a,b)\delta + G_b(a,b)\epsilon$$

josta saadaan yhtälöpari 'korjaukselle' (δ, ϵ) :

$$(7) \quad F_a \cdot \delta + F_b \cdot \epsilon = -F$$

$$(8) \quad G_a \cdot \delta + G_b \cdot \epsilon = -G .$$

Tässä osittaisderivaatat ovat

$$(10) \quad F_a(a,b) = \psi'(a+b) - \psi'(a)$$

$$(11) \quad F_b(a,b) = \psi'(a+b)$$

$$(12) \quad G_a(a,b) = \psi'(a+b)$$

$$(13) \quad G_b(a,b) = \psi'(a+b) - \psi'(b) \quad .$$

Käyttämällä matriisimerkintöjä

$$(14) \quad Y1 = \begin{bmatrix} \delta \\ \epsilon \end{bmatrix}, \quad H = \begin{bmatrix} F_a & F_b \\ G_a & G_b \end{bmatrix}, \quad Y = \begin{bmatrix} -F \\ -G \end{bmatrix}$$

saadaan yhtälöt (8) - (9) ratkaistua muodossa

$$(15) \quad Y1 = H^{-1}Y \quad .$$

Varsinainen algoritmi on muotoa

$$(16) \quad \begin{bmatrix} a \\ b \end{bmatrix}_{n+1} = \begin{bmatrix} a \\ b \end{bmatrix}_n + \begin{bmatrix} \delta \\ \epsilon \end{bmatrix}_n \quad .$$

Oheisessa tietokoneohjelmassa on sovellettu psii-funktion ja sen derivaatan approksimaatiota

$$(17) \quad \psi(x) \approx \log x - \frac{A(x)}{2x}, \quad \text{jossa}$$

$$(18) \quad A(x) = 1 + \frac{1}{6x + \frac{6}{10x + (\frac{6}{1 + 1.2879})}}$$

$$(19) \quad \psi'(x) \approx \frac{1}{x - \frac{1}{2} + \frac{1}{12x + 10}} .$$

Ensin on tosin käytetty funktioiden siirto-ominaisuuksia

$$(20) \quad \psi(x) = \psi(x+1) - \frac{1}{x}$$

$$(21) \quad \psi'(x) = \psi'(x+1) + \frac{1}{x^2} .$$

Algoritmi konvergoi kaikkialla nopeasti ainakin jos alkuarvot poikkeavat ratkaisusta alle 10 %. Tietokoneohjelma on kirjoitettu GE 600 BASIC-kielellä. Ohjelman jäljessä on esimerkkiajo, joka ratkaisee arvoihin $u = 36.3093$, $v = 0.4748$ liittyvät ml-yhtälöt, alkuarvoina on käytetty $a_0 = 0.03$, $b_0 = 0.3$.

```

10 REM THE PROGRAM CALCULATES THE ML ESTIMATES OF BETA DISTRIBUTION
20 REM YRJO.O.VARTIA 10.7.1973
25 DIM Y(2,1),H(2,2),K(2,2)
26 DIM Y1(2,1)
30 PRINT "CHARACTERISTICS U,V";
35 PRINT
40 INPUT U,V
50 DEF APU(X)=1+1/(6*X+6/(10*X+6/(1+1.2879*X)))
55 DEF FNA(X)=LOG(X)-APU(X)/(2*X)
60 DEF PSI(X)=FNA(X+1)-1/X
70 DEF FNB(X)=1/(X+.5+1/(12*X+10))+1/X^2
80 PRINT "FIRST ESTIMATES OF A,B";
82 PRINT
85 INPUT A,B
91 I=0
92 I=I+1
95 PRINT
100 X=A+B
110 Y(1,1)=U+PSI(A)-PSI(X)
120 Y(2,1)=V+PSI(B)-PSI(X)
130 H(1,1)=FNB(X)-FNB(A)
140 H(1,2)=FNB(X)
150 H(2,1)=H(1,2)
160 H(2,2)=FNB(X)-FNB(B)
170 MATK=INV(H)
175 MAT Y1=K*Y
190 A=A+Y1(1,1)
200 B=B+Y1(2,1)
210 PRINT USING 299,A,B
220 IF I<6 THEN 92
225 PRINT
230 PRINT"STOP ? NO=0";
240 INPUT S
250 IF S=0 THEN 91
299:ML-ESTIMATES      A = ###.#####      B = ###.#####
300 END

```

CHARACTERISTICS U,V

? 36.3093, 0.4748

FIRST ESTIMATES OF A,B

? .03, .3

ML-ESTIMATES A = .02396 B = .19793

ML-ESTIMATES A = .02495 B = .22117

ML-ESTIMATES A = .02500 B = .22491

ML-ESTIMATES A = .02500 B = .22500

ML-ESTIMATES A = .02500 B = .22500

ML-ESTIMATES A = .02500 B = .22500

STOP ? NO=0 ? 1

LIITE 4. LIKELIHOOD YHTÄLÖIDEN RATKAISUJA $\theta = 0.1, 0.2, \dots, 0.9$ JA $p = 0.1, 0.2, \dots, 0.5$ VASTAAVAT HAVAINTOARVOT u JA v SEKÄ f JA g

θ	p	a	b	u	v	$f = e^{-u}$	$g = e^{-v}$
0.1	0.1	0.9	8.1	2.8956	0.1118	0.05527	0.8942
	.2	1.8	7.2	1.8557	0.2376	0.1564	0.7885
	.3	2.7	6.3	1.3439	0.3816	0.2608	0.6828
	.4	3.6	5.4	1.0050	0.5497	0.3661	0.5771
	.5	4.5	4.5	0.7518	0.7518	0.4715	0.4715
0.2	0.1	0.4	3.6	3.8175	0.1205	0.02198	0.8865
	.2	0.8	3.2	2.2211	0.2573	0.1085	0.7732
	.3	1.2	2.8	1.5452	0.4156	0.2133	0.6600
	.4	1.6	2.4	1.1301	0.6032	0.3230	0.5470
	.5	2.0	2.0	0.8333	0.8333	0.4346	0.4346
0.3	0.1	0.2333	2.1000	5.1513	0.1326	0.00579	0.8758
	.2	0.4667	1.8667	2.7560	0.2850	0.06355	0.7520
	.3	0.7000	1.6333	1.8380	0.4637	0.1591	0.6290
	.4	0.9333	1.4000	1.3105	0.6794	0.2697	0.5069
	.5	1.1667	1.1667	0.9501	0.9501	0.3867	0.3867
0.4	0.1	0.15	1.35	7.0575	0.1504	$860 \cdot 10^{-6}$	0.8603
	.2	0.30	1.20	3.5390	0.3255	0.02904	0.7221
	.3	0.45	1.05	2.2700	0.5343	0.1033	0.5861
	.4	0.60	0.90	1.5771	0.7914	0.2066	0.5432
	.5	0.75	0.75	1.1224	1.1223	0.3255	0.3255
0.5	0.1	0.10	0.90	9.8465	0.1777	$529 \cdot 10^{-7}$	0.8372
	.2	0.20	0.80	4.7118	0.3878	0.00899	0.6786
	.3	0.30	0.70	2.9253	0.6428	0.05365	0.5258
	.4	0.40	0.60	1.9842	0.9634	0.1375	0.3816
	.5	0.50	0.50	1.3863	1.3863	0.2500	0.2500
0.6	0.1	0.0667	0.6000	14.1544	0.2224	$713 \cdot 10^{-9}$	0.8006
	.2	0.1333	0.5333	6.5588	0.4896	0.00142	0.6129
	.3	0.2000	0.4667	3.9708	0.8198	0.01886	0.4405
	.4	0.2667	0.4000	2.6394	1.2432	0.07140	0.2885
	.5	0.3333	0.3333	1.8138	1.8138	0.16303	0.1630
0.7	0.1	0.0429	0.3857	21.4764	0.3029	$471 \cdot 10^{-12}$	0.7387
	.2	0.0857	0.3429	9.7453	0.6725	$586 \cdot 10^{-7}$	0.5104
	.3	0.1286	0.3000	5.7955	1.1367	0.00304	0.3209
	.4	0.1714	0.2571	3.7934	1.7420	0.02252	0.1752
	.5	0.2143	0.2143	2.5719	2.5719	0.07639	0.07639
0.8	0.1	0.0250	0.2250	36.3094	0.4748	$170 \cdot 10^{-18}$	0.6220
	.2	0.0500	0.2000	16.2704	1.0616	$859 \cdot 10^{-10}$	0.3459
	.3	0.0750	0.1750	9.5661	1.8080	$701 \cdot 10^{-7}$	0.1640
	.4	0.1000	0.1500	6.1963	2.7935	$204 \cdot 10^{-5}$	0.06120
	.5	0.1250	0.1250	4.1610	4.1610	0.01559	0.01559
0.9	0.1	0.0111	0.1000	81.1511	1.0158	$571 \cdot 10^{-38}$	0.36211
	.2	0.0222	0.0889	36.1333	2.2819	$203 \cdot 10^{-18}$	0.10209
	.3	0.0333	0.0778	21.1157	3.9053	$675 \cdot 10^{-12}$	0.02014
	.4	0.0444	0.0667	13.5984	6.0646	$124 \cdot 10^{-8}$	0.00232
	.5	0.0555	0.0555	9.0814	9.0814	0.00011	0.00011

LIITE 5. JOITAKIN TÄRKEITÄ VAKIOITA

γ	$= -\psi(1)$	$=$	0.577	215	664	901	533
e^γ		$=$	1.781	072			
$e^{-\gamma}$		$=$	0.561	460			
	$\psi^{-1}(0)$	$=$	1.461	632	144	968	362
	$\psi^{-1}(-1)$	$=$	0.785	00			
	$\psi^{-1}(-2)$	$=$	0.492	70			
	$\psi^{-1}(-3)$	$=$	0.346	89			
$\frac{\pi^2}{6}$	$= \zeta(2)$	$=$	1.644	934	067		
	$\zeta(3)$	$=$	1.202	056	903		
$\frac{\pi^4}{90}$	$= \zeta(4)$	$=$	1.082	323	234		
	$\zeta(5)$	$=$	1.036	927	755		
	$\zeta(6)$	$=$	1.017	343	062		
π^2		$=$	9.869	604	401	089	358
$\frac{\pi^2}{4}$		$=$	2.467	401	100	272	339
$-\gamma - \log 4$	$= \psi\left(\frac{1}{2}\right)$	$=$	-1.963	510	026	021	423

LIITE 6. ARVIOIDEN $1/x \approx 2d$ JA $p \approx \frac{f+d}{1+d}$ TARKKUUDET

θ	$\frac{1}{x} = \frac{\theta}{1-\theta}$	p	$2d$	$\frac{f+d}{1+d}$
0.1	0.1111	0.1	0.1010	0.1007
		.2	.1103	.2004
		.3	.1127	.3003
		.4	.1136	.4001
		.5	.1139	.5000
0.2	0.25	0.1	0.1830	0.1040
		.2	.2367	.2028
		.3	.2535	.3018
		.4	.2599	.4009
		.5	.2616	.5000
0.3	0.4286	0.1	0.2368	0.1111
		.2	.3688	.2094
		.3	.4238	.3062
		.4	.4468	.4030
		.5	.4532	.5000
0.4	0.6667	0.1	0.2776	0.1226
		.2	.4976	.2225
		.3	.6213	.3158
		.4	.6805	.4080
		.5	.6979	.5000
0.5	1.0	0.1	0.3255	0.1400
		.2	.6249	.2449
		.3	.8411	.3338
		.4	.9618	.4176
		.5	1.0000	.5000

LIITE 7. ARVIOIDEN $1/x \approx \sqrt{uv}$ ja $p \approx 1/(1 + \sqrt{\frac{u}{v}})$ TARKKUUDET

θ	$\frac{1}{x} = \frac{\theta}{1-\theta}$	p	\sqrt{uv}	$\frac{1}{1 + \sqrt{\frac{u}{v}}}$
0.5	1.0	0.1	1.3228	.1184
		.2	1.3517	.2229
		.3	1.3713	.3192
		.4	1.3826	.4107
		.5	1.3863	.5000
0.6	1.5	0.1	1.7742	.1114
		.2	1.7919	.2146
		.3	1.8042	.3124
		.4	1.8114	.4070
		.5	1.8138	.5000
0.7	2.3333	0.1	2.5506	.1062
		.2	2.5600	.2080
		.3	2.5667	.3069
		.4	2.5706	.4039
		.5	2.5719	.5000
0.8	4.0	0.1	4.1521	.1026
		.2	4.1560	.2035
		.3	4.1588	.3030
		.4	4.1605	.4017
		.5	4.1610	.5000
0.9	9.0	0.1	9.0793	.10062
		.2	9.0802	.20083
		.3	9.0809	.30073
		.4	9.0813	.40041
		.5	9.0814	.50000

LIITE 8. ARVIOIDEN z_0 JA z_2 TARKKUUDET

θ	p	$z = e^{\psi(a+b)}$	z_0	z_2
0.1	0.1	8.5049	9.0316	8.2665
	.2		8.5099	8.4900
	.3		8.3926	8.5017
	.4		8.3533	8.5036
	.5		8.3431	8.5040
0.2	0.1	3.5118	3.8925	4.6062
	.2		3.5767	3.3571
	.3		3.4539	3.4813
	.4		3.4088	3.5004
	.5		3.3969	3.5039
0.3	0.1	1.8552	1.8632	3.7304
	.2		1.9023	2.0148
	.3		1.8372	1.7000
	.4		1.8022	1.8009
	.5		1.7919	1.8176
0.4	0.1	1.0372	1.0020	2.4135
	.2		1.0393	1.7885
	.3		1.0345	1.2474
	.4		1.0214	0.9081
	.5		1.0164	0.8217
0.5	0.1	0.56146	0.5658	2.2993
	.2		0.5600	1.3068
	.3		0.5646	1.0572
	.4		0.5644	0.8919
	.5		0.5637	0.8361
0.6	0.1	0.26761	0.2837	
	.2		0.2781	
	.3		0.2775	
	.4		0.2781	
	.5		0.2784	
0.7	0.1	0.09387	0.1062	
	.2		0.1050	
	.3		0.1044	
	.4		0.1042	
	.5		0.1042	
0.8	0.1	0.01459	0.01838	
	.2		0.01830	
	.3		0.01825	
	.4		0.01822	
	.5		0.01821	
0.9	0.1	$8.207 \cdot 10^{-5}$	$12.17 \cdot 10^{-5}$	
	.2		$12.16 \cdot 10^{-5}$	
	.3		$12.15 \cdot 10^{-5}$	
	.4		$12.15 \cdot 10^{-5}$	
	.5		$12.15 \cdot 10^{-5}$	

Arviot z_0 ja z_2 on esitetty sivulla 70 ja 71.

LIITE 9. ARVIOIDEN a_2 JA b_2 TARKKUUDET

θ	p	a	b	a_2	b_2
0.1	0.1	0.9	8.1	0.8758	7.8867
	.2	1.8	7.2	1.7969	7.1882
	.3	2.7	6.3	2.6990	6.2978
	.4	3.6	5.4	3.5995	5.3993
	.5	4.5	4.5	4.4995	4.4995
0.2	0.1	0.4	3.6	0.3658	4.5733
	.2	0.8	3.2	0.7662	3.0796
	.3	1.2	2.8	1.1898	2.7796
	.4	1.6	2.4	1.5950	2.3935
	.5	2.0	2.0	1.9960	1.9960
0.3	0.1	0.2333	2.1000	0.1178	3.7544
	.2	0.4667	1.8667	0.4217	1.9883
	.3	0.7000	1.6333	0.6466	1.5316
	.4	0.9333	1.4000	0.9088	1.3694
	.5	1.1667	1.1667	1.1475	1.1475
0.4	0.1	0.15	1.35	0.0133	2.5566
	.2	0.30	1.20	0.2343	1.7602
	.3	0.45	1.05	0.4233	1.1776
	.4	0.60	0.90	0.5259	0.8230
	.5	0.75	0.75	0.6424	0.6424
0.5	0.1	0.10	0.90	0.0008	2.4036
	.2	0.20	0.80	0.0690	1.3420
	.3	0.30	0.70	0.2494	0.9876
	.4	0.40	0.60	0.4110	0.7368
	.5	0.50	0.50	0.5591	0.5591
0.6	0.1	0.0667	0.6000		
	.2	0.1333	0.5333		
	.3	0.2000	0.4667		
	.4	0.2667	0.4000		
	.5	0.3333	0.3333		
0.7	0.1	0.0429	0.3857		
	.2	0.0857	0.3429		
	.3	0.1286	0.3000		
	.4	0.1714	0.2571		
	.5	0.2143	0.2143		
0.8	0.1	0.0250	0.2250		
	.2	0.0500	0.2000		
	.3	0.0750	0.1750		
	.4	0.1000	0.1500		
	.5	0.1250	0.1250		
0.9	0.1	0.0111	0.1000		
	.2	0.0222	0.0889		
	.3	0.0333	0.0778		
	.4	0.0444	0.0667		
	.5	0.0555	0.0555		

Arviot a_2 ja b_2 on esitetty sivulla 71.

LIITE 10. ARVIOIDEN a_3 JA b_3 TARKKUUDET

θ	p	a	b	a_3	b_3
0.1	0.1	0.9	8.1	0.9275	8.4626
	.2	1.8	7.2	1.8030	7.2188
	.3	2.7	6.3	2.6787	6.2478
	.4	3.6	5.4	3.5576	5.3346
	.5	4.5	4.5	4.4415	4.4415
0.2	0.1	0.4	3.6	0.4159	3.8283
	.2	0.8	3.2	0.8031	3.2345
	.3	1.2	2.8	1.1856	2.7699
	.4	1.6	2.4	1.5693	2.3541
	.5	2.0	2.0	1.9573	1.9573
0.3	0.1	0.2333	2.1000	0.2435	2.1448
	.2	0.4667	1.8667	0.4707	1.8881
	.3	0.7000	1.6333	0.6926	1.6186
	.4	0.9333	1.4000	0.9150	1.3729
	.5	1.1667	1.1667	1.1405	1.1405
0.4	0.1	0.15	1.35	0.1544	1.3524
	.2	0.30	1.20	0.3090	1.2113
	.3	0.45	1.05	0.4507	1.0447
	.4	0.60	0.90	0.5931	0.8877
	.5	0.75	0.75	0.7382	0.7382
0.5	0.1	0.10	0.90	0.1017	0.9031
	.2	0.20	0.80	0.2093	0.8115
	.3	0.30	0.70	0.3101	0.7066
	.4	0.40	0.60	0.4049	0.6012
	.5	0.50	0.50	0.5012	0.5012
0.6	0.1	0.0667	0.6000	0.0673	0.6062
	.2	0.1333	0.5333	0.1378	0.5450
	.3	0.2000	0.4667	0.2116	0.4831
	.4	0.2667	0.4000	0.2821	0.4164
	.5	0.3333	0.3333	0.3491	0.3491
0.7	0.1	0.0429	0.3857	0.04315	0.39989
	.2	0.0857	0.3429	0.08754	0.36019
	.3	0.1286	0.3000	0.13429	0.32047
	.4	0.1714	0.2571	0.18361	0.27875
	.5	0.2143	0.2143	0.23293	0.23293
0.8	0.1	0.0250	0.2250	0.02506	0.23577
	.2	0.0500	0.2000	0.05026	0.20856
	.3	0.0750	0.1750	0.07561	0.18109
	.4	0.1000	0.1500	0.10120	0.15385
	.5	0.1250	0.1250	0.12721	0.12721
0.9	0.1	0.0111	0.1000	0.01108	0.09914
	.2	0.0222	0.0889	0.02212	0.08807
	.3	0.0333	0.0778	0.03312	0.07704
	.4	0.0444	0.0667	0.04410	0.06605
	.5	0.0555	0.0555	0.05507	0.05507

Arviot a_3 ja b_3 on esitetty sivulla 77.

LIITE 11. ARVIOIDEN a_4 JA b_4 TARKKUUDET

θ	p	a	b	a_4	b_4
0.1	0.1	0.9	8.1	0.8999	8.0911
	.2	1.8	7.2	1.7757	7.0750
	.3	2.7	6.3	2.5303	5.8547
	.4	3.6	5.4	2.9616	4.3929
	.5	4.5	4.5	2.8429	2.8407
0.2	0.1	0.4	3.6	0.3988	3.5894
	.2	0.8	3.2	0.7982	3.1082
	.3	1.2	2.8	1.1715	2.7081
	.4	1.6	2.4	1.4596	2.1613
	.5	2.0	2.0	1.5482	1.5484
0.3	0.1	0.2333	2.1000	0.2321	2.0969
	.2	0.4667	1.8667	0.4672	1.8670
	.3	0.7000	1.6333	0.6955	1.6105
	.4	0.9333	1.4000	0.8883	1.3173
	.5	1.1667	1.1667	0.9764	0.9849
0.4	0.1	0.15	1.35	0.1496	1.3638
	.2	0.30	1.20	0.3013	1.2165
	.3	0.45	1.05	0.4527	1.0591
	.4	0.60	0.90	0.5866	0.8808
	.5	0.75	0.75	0.6499	0.6769
0.5	0.1	0.10	0.90		
	.2	0.20	0.80		
	.3	0.30	0.70		
	.4	0.40	0.60		
	.5	0.50	0.50		
0.6	0.1	0.0667	0.6000		
	.2	0.1333	0.5333		
	.3	0.2000	0.4667		
	.4	0.2667	0.4000		
	.5	0.3333	0.3333		
0.7	0.1	0.0429	0.3857		
	.2	0.0857	0.3429		
	.3	0.1286	0.3000		
	.4	0.1714	0.2571		
	.5	0.2143	0.2143		
0.8	0.1	0.0250	0.2250		
	.2	0.0500	0.2000		
	.3	0.0750	0.1750		
	.4	0.1000	0.1500		
	.5	0.1250	0.1250		
0.9	0.1	0.0111	0.1000		
	.2	0.0222	0.0889		
	.3	0.0333	0.0778		
	.4	0.0444	0.0667		
	.5	0.0555	0.0555		

Arviot a_4 ja b_4 on esitetty sivulla 83 ja 84.

PENTTI LAURI ILMARI VARTIA

HUOMIOITA TEOREETTISTEN APUKÄSITTEIDEN IDENTIFIOITUVUUDESTA

COMMENTS ON THE IDENTIFICATION OF AUXILIARY CONCEPTS

Helsinki 1973

SISÄLLYSLUETTELO

	sivu
0. JOHDANTO	5
1. IDENTIFIOITUVUUDESTA YLEISESTI	7
2. IDENTIFIOITAVUUDESTA LINEAARISISSA MALLEISSA	16
2.1. A priori malli	16
2.2. A posteriori observoimalla saatava tieto	19
2.3. A priori tiedon ja a posteriori tiedon yhdistäminen ja parametrien identifioiminen	21
2.4. Rakennemuodon yhtälöiden lineaarikombinaatiosta	23
2.5. Identifioituvuus ja redusoitu muoto	24
3. ERÄÄSTÄ TERMIN ALI-IDENTIFIOITUVUUS KÄYTÖSTÄ	27
4. YLI-IDENTIFIOITUVUUDESTA	29
5. HUOMIOITA IDENTIFIOITUVUUDESTA JA MÄÄRITELTÄVYYDESTÄ	33
6. HUOMIOITA TEOREETTISTEN APUKÄSITTEIDEN TARPEELLISUUDESTA	36
7. LOPPUPÄÄTELMIÄ	44
KIRJALLISUUSVIITTEET	46

0. JOHDANTO

Identifiointiongelmaan on ekonometrisessa kirjallisuudessa kiinnitetty runsaasti huomiota, mm. melkein kaikki standardioppikirjat omistavat luvun ongelman esittämiselle. Aihepiirin käsittely on kuitenkin melko epäyhtenäistä, esimerkiksi käytetty terminologia eroaa eri kirjoittajilla. Sen yksinkertaisen asian esittäminen, että samat ilmiöt voidaan selittää monella eri tavalla, näyttää joskus vaativan kovin paljon ja kovin mutkikasta aherrusta, jonka alle identifiointiongelmaan liittyvät loogiset kysymykset helposti peittyvät. Tässä esitettävä tarkastelu on pyritty suorittamaan niin, että identifiointiongelman yhteydet määriteltävyyden käsitteeseen olisivat nähtävissä. Tällöin terminologiset vaikeudet yhä korostuvat sillä määriteltävyyttä on tutkittu lähinnä logiikan puolella ja näin käytetyt termit eroavat ekonometrisessä kirjallisuudessa esiintyvistä. Koska lähtökohtana on kuitenkin juuri identifiointiongelma on työssä pitäydytty jonkinlaiseen ekonometrisen kirjallisuuden keskimääräiseen terminologiaan. Sekaanusta aiheuttaa myös se, että identifiointia voidaan tutkia deterministisissä ja stokastisissa malleissa, sekä stokastisessa tapauksessa vielä populaatio- tai otostasolla. Havaintojen lukumäärän vaikutus parametrien "identifiointiin" on eri tapauksissa erilainen.

Esimerkkinä identifioitavuuden esittelyssä on käytetty lineaarista ekonometrista mallia, jonka ominaisuuksia mm. identifioitavuutta on kirjallisuudessa melko perusteellisesti tutkittu. Pyrkimyksenä on esitellä ongelmaa hiukan laajemmin ja välttää liiallista teknillistä keikarointia, joka saattaa peittää oleellisesti yksinkertaisetkin kysymyksenasettelut.

Haluan lausua kiitokseni prof. Raimo Tuomelalle ja VTK, LuK Yrjö Vartialle saamistani huomautuksista.

1. IDENTIFIOITAVUUDESTA YLEISESTI

Ekonometrissa mallia (vrt. logiikan teoria) on seuraavassa karakterisoitu rakenteiden (vrt. logiikan malli) joukkona, jotka ovat sopusoinnussa tiettyjen aksioomien kanssa. Mallilla ymmärretään muuttujien välisten relaatioiden muodon ja virhetermin jakauman muodon spesifioimista. Kun mallin parametreille (vrt. teoreettiset apukäsitteet) annetaan tietyt numeroarvot saadaan eräs malliin kuuluva rakenne.

Koska ekonometriset mallit pyrkivät kuvailemaan empiirisiä ilmiöitä, on mallia muodostettaessa ajateltava ainakin osa mallin muuttujista havaittavissa oleviksi ja niiden tulkin-
ta kiinnitetyksi. Tosin useissa ekonometrisissa malleissa monet "observaatiotermit" ovat etäänntyneet melko kauas esim. ostensiivisesta määrittelyn mahdollisuudesta, pikemminkin meillä on toisia teorioita, joiden avulla ko. käsitteet tulevat määriteltyä. Em. problematiikkaan ei tässä liene syytä puuttua, oletettakoon vain yksinkertaisesti että meillä on jokin tapa millä observaatiotermit on saatu määriteltyä.

Mallilla pyritään esittämään kaikki a priori informaatio mekanismista, jolla oletamme havaintojen syntyvän. Tällöin aksioomiin on luettava mukaan myös kaikki mallin parametreja

koskevat a priori rajoitukset. Termin a priori katsotaan tässä liittyvän juuri ko. mallinrakennustilanteeseen. Kyseessä saattaa olla paitsi teoreettisten pohdiskelujen, myös empiiristen kokemusten kautta muualta tai aikaisemmin saatu tietämys. Usein lisäksi mallin empiirinen soveltaminen antaa aihetta mallin pieniin muutoksiin, jolloin a priori ja a posteriori tietoa on vaikea erottaa. Yksinkertaisinta on tällöin olettaa, että joka kerta on kyse aivan uudesta mallista.

Huomattakoon, että todellisuudessa mallinrakentajan ei useinkaan olisi syytä uskoa malliinsa kovin pitkälle. Sen tulisi olla hänelle paremminkin eräänlainen "leikkiprosessi", jota hän käyttää "reaaliprosessin" kuvaamiseen.

Tarkasteltaessa identifiointiongelmia stokastisissa malleissa oletetaan usein, että tuntemme riippuvien muuttujien ehdollisen jakauman (ehdolla, että ennalta määrätyt muuttujat saavat tietyt arvot). Tällaista tietoa ei luonnollisestikaan voida saavuttaa äärellisellä otoksella. "Äärettömän otoksen" käyttö on perusteltua, koska identifiointiongelmia eroaa selvästi estimointiongelmista, loogisesti edeltäen niitä ja tällä oletuksella ero saadaan selvästi esille. Identifiointi asettaa nimittäin selvät rajat tilastolliselle päättelylle, vaikka olisi käytettävissä äärettömiäkin otoksia.

Mallin riippuvien, endogeenisten muuttujien jakauma jota mallilla pyritään selittämään, riippuu oleellisesti muiden, ennaltamäärättyjen muuttujien arvoista (jotka seuraavassa oletetaan kiinteiksi luvuiksi) ja häiriötermin jakaumasta. Ennalta määrättyt muuttujat ja ainakin osa riippuvista muuttujista on havaittavissa, sen sijaan malliin liittyvää häiriötermiä ei voida suoraan havainnoida.

Sen havaintojen generoimistavan, jonka mallimme kulloinkin kiinnittää, esittäminen yleisesti on usein hankalaa ja epähavainnollista. Eräs mahdollinen, nimenomaan stokastisiin malleihin sopiva yleinen tapa on ajatella riippuvien muuttujien y jakauman G muodostuvan häiriötermin u jakaumasta F operaation I kautta, eli¹⁾

$$G = (I \rightarrow F) \quad (1.1.)$$

Tavallisesti operaatioon I vaikuttavat ennaltamäärättyt muuttujat z . Oleellista on että se sisältää tutkijalle tuntemattomia teoreettisia apukäsitteitä, parametreja.

1) Tässä esitetty mallin määritelmä seuraa Hurwicz (1):tä. Identifioituvuus on määritelty hiukan eri lailla, koska tuntuu luonnolliselta mainita relaation R yksikäsitteisyys määritelmässä. n -ali-identifioituvuuden määritelmä seuraa Tuomela (1):tä. Kuitenkin on haluttu erottaa n -ali-identifioituvuus voimakkaammin identifioituvuudesta vaatimalla etteivät n -ali-identifioituvat mallit voi olla identifioituvia, mikä on mahdollista Tuomela (1):ssä.

Usein on helpompi käsitellä satunnaismuuttujien, kuin niiden jakaumien välisiä operaatioita. Jos esim. vektori \underline{y} saadaan \underline{u} :sta ja z :sta muunnoksen T avulla seuraavasti

$$\underline{y} = T(z, \underline{u}) \quad (1.2.)$$

määräytyy (tietyillä ehdoilla) myös \underline{y} :n jakauma yhtälöstä

$$P\{\underline{y} \leq y\} = P\{\underline{y} \mid T(z, \underline{u}) \leq y\} \quad (1.3.)$$

Muunnos T määrää näin ollen jakaumien välisen operaation. T :n muoto lineaarisissa malleissa on esitetty luvussa 2.1. Jakauman F muoto ja operaatio I (sisältää vapaita parametreja) sisältävät siis informaation mekanismista, jolla havaintojen postuloidaan syntyneen. Tiettyä paria F ja I (jossa F ja I kiinteitä) vastaa tietty mallin parametrien kiinnittäminen, eli tietty mallin rakenne. Kutsumme kaksikkoa

$$S = (F, I)$$

rakenteeksi S . Mikäli (1.1.) on voimassa sanotaan, että rakenne S generoi jakauman G , ja merkitään $G=R(S)$. R on siis funktio, jonka määrittelyalue on rakenteiden joukko \mathcal{S} ja maalialue endogeenisten muuttujien jakaumien G joukko \mathcal{E} . Se, että R on funktio, johtuu operaation I yksikäsitteisyydestä. Jokaista rakennetta S vastaa siis yksikäsitteinen jakauma G , eli

$$\forall S, \forall G_1, \forall G_2 \left[G_1 = R(S) \ \& \ G_2 = R(S) \rightarrow G_1 = G_2 \right]$$

Muodostamme käyttämämme mallin valitsemalla a priori kaikkien mahdollisten rakenteiden joukosta tietyn osajoukon, eli rajoittamalla niitä tapoja, joilla a priori olemme valmiit hyväksymään havainnot generoiduiksi. Tällä tavoin karsimme siis pois selitykset, jonka tyyppisiäkään emme periaatteessa hyväksy. Havainnot voidaan selittää monella tavalla, mutta niiden katseleminen ei sinänsä useinkaan riitä asettamaan kilpailevia selityksiä toistensa edelle. Emme yksinkertaisesti voi hyväksyä kaikenlaisia selityksiä, esimerkiksi sellaisia, joissa kotitonttu sopivassa mielentilassa kriittisellä hetkellä puuttuu asioiden kulkuun. Muodollisesti voimme siis määritellä mallimme seuraavasti:

MÄÄRITELMÄ: Malli M on a priori valittu kaikkien mahdollisten rakenteiden S joukon \mathcal{S} osajoukko M .

Tarkastellaan niiden jakaumien G joukkoa K , jotka voidaan generoida malliimme M kuuluvilla rakenteilla S

$$K = \{ G \mid G = R(S) \ \& \ S \in M \} \quad (1.5.)$$

Mikäli nyt jokaiselle K :n alkiolle G pitää paikkansa, että sitä vastaa joukossa M ainoastaan yksi alkio S , eli

$$\forall G \{ G \in K \Rightarrow [(G = R(S_1) \ \& \ G = R(S_2)) \Rightarrow S_1 = S_2] \} \quad (1.6.)$$

katakin endogeenisten muuttujien jakaumaa vastaava rakenne on identifioitavissa. Identifioituva malli voidaankin määritellä seuraavasti:

MÄÄRITELMÄ: Malli M on identifioituva, mikäli relaatio R rakenteiden S ja jakaumien G välillä on kääntäen yksikäsitteinen M:ssä.

Mikäli malli ei ole identifioituva, se on ei-identifioituva. Usein mallia tällöin kutsutaan ali-identifioituvaksi, mikä mielestäni on hiukan harhaanjohtava termi. Tätä termiä käytetään kirjallisuudessa myös hiukan poikkeavassa mielessä (ks. luku 3).

Vaikka malli ei olisikaan identifioituva, voi joku sen osa olla tätä, esim. yksityinen parametri on identifioituva mikäli sillä on sama arvo kaikissa niissä mallin M rakenteissa, jotka generoivat saman endogeenisten muuttujien jakauman. Mallin tiettyä osaa, esim. yhtälöä kutsutaan identifioituvaksi, mikäli kaikki siihen liittyvät parametrit ovat identifioituvia (eräissä tapauksissa lienee kuitenkin epäselvää, milloinka tietyn parametrin on katsottava liittyvän tiettyyn yhtälöön, milloin ei).

Mallin M identifioitavuus merkitsee siis sitä, että kutakin observoitavaa endogeenisten muuttujien jakaumaa vastaa M:ssä yksikäsitteinen rakenne. M:ssä ei ole olemassa useampia observationaalisesti ekvivalentteja rakenteita. Mallin osan identifioitavuus taas merkitsee sitä että kaikissa tietyn jakauman määrittelemissä (rakenteiden) ekvivalenssiluokkien rakenteissa tähän mallin osaan liittyvien parametrien arvot ovat samat.

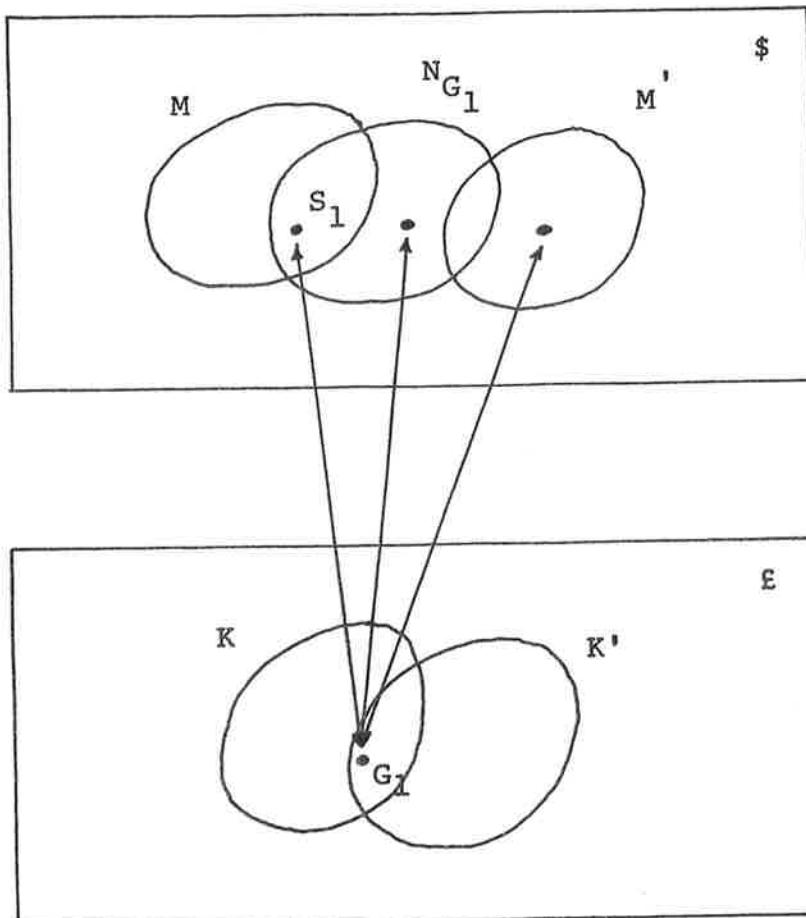
Identifioitavuuden käsite voidaan laajentaa normaalin identifioitavuuden ulkopuolelle tarkastelemalla ei-identifioituvia malleja tai mallin osia ja kiinnittämällä huomio siihen kuinka monta $M:n$ rakennetta kuuluu tietyn jakauman määräämään rakenteiden ekvivalenssiluokkaan, tai kuinka monta mahdollista arvoa tiettyyn mallin osaan liittyvät parametrit voivat tällaisessa ekvivalenssiluokassa saada.

MÄÄRITELMÄ: Malli M on n -ali-identifioituva, jos kaikille sen rakenteiden generoimille endogeenisten muuttujien jakaumille $G \in K$ pitää paikkansa, että niiden määräämiin rakenteiden ekvivalenssiluokkiin kuuluu korkeintaan n rakennetta $M:ssä$ ja ainakin yhteen ekvivalenssiluokkaan kuuluu täsmälleen n rakennetta $M:ssä$.

Edellä esitetty n -ali-identifioitavuus (n äärellinen) on yleistettävissä numeroituvaan ja ylinumeroituvaan identifioitavuuden käsitteeseen. Samoin se voidaan laajentaa koskemaan tiettyä mallin osaa.

Huomautettakoon, että aikaisemmin käytetty termi n -identifioituva on mielestäni harhaanjohtava, ja olisi korvattava toisella. Sanallisesti yhteys n -identifioitavuuden ja "yli-identifioitavuuden" välillä on liian läheinen, vaikka termit pyrkivät kuvaamaan kvalitatiivisesti kahta erilaista ilmiötä. Mahdollinen olisi esimerkiksi termi n -kertaisesti ali-identifioituva tai edellä esitetty n -ali-identifioituva.

Valitessamme a priori malliamme M voimme tehdä siitä laajan tai rajoittavan. Voimme esim. tehdä siitä niin rajoittavan, että kun observaatiotermien tulkinta on kiinnitetty, ei mikään mallimme rakenteista ole sopusoinnussa havaintojen kanssa, eli leikkiprosessimme ei sovi kuvailemaan reaaliprosessia, kunhan meillä vaan on tarpeeksi havaintoja tämän toteamiseen. Ihanteellinen tapaus on luonnollisesti sellainen, jossa mallistamme löytyy täsmälleen yksi rakenne, joka on sopusoinnussa havaintojen kanssa. Mallin identifioitavuus takaa juuri sen, että mikäli mallissamme on jokin havaintojen kanssa sopusoinnussa oleva rakenne S_1 , niin samojen havaintojen kanssa sopusoinnussa olevaa toista rakennetta ei M :stä löydy. Tilannetta on tarkasteltu oheisessa kuviossa.



Havaintojen perusteella (havainnot ymmärrettynä riittävän laajasti) voidaan kiinnittää tietty endogeenisten muuttujien jakauma G_1 . Tämä määrää ko. jakaumaa vastaavan ekvivalenssi-luokan N_{G_1} eli niiden rakenteiden luokan jotka ovat sopu-soinnussa havaintojen kanssa (data admissible set of structures, Christin terminologian mukaan).

Normaalissa tilanteessa havaintojen kanssa muita sopusoinnussa olevia rakenteita löytyy kyllä mallin ulkopuolelta, esim. jostain toisesta mallista M' . Samalle ilmiölle voidaan siis käyttää toista mallia, toisenlaisia a priori rajoituksia ja oletuksia muuttujien välisten relaatioiden muodosta (esim. epälineaarista mallia lineaarisen asemasta). Mikäli rajoitumme vain lineaarisiin malleihin (emme tee erityis-oletuksia virhetermeistä) N tulee vastaamaan niiden mallien joukkoa, joilla on sama redusoitu muoto.

2. IDENTIFIOITAVUUDESTA LINEAARISISSA MALLEISSA

2.1. A priori malli

Edellä on yleisin termein kuvattu sitä, mitä identifioitavuus merkitsee puuttumatta siihen miten yksityisessä tilanteessa voidaan todeta onko jokin malli tai sen osa identifioituva vai ei, tai miten itse identifioiminen sitten suoritetaan.

Mallista riippuen on usein mahdollista tutkia edellytyksiä, joiden vallitessa tiettyyn malliryhmään kuuluvat mallit ovat identifioituvia. Täten voidaan päätyä erilaisiin sääntöihin jotka takaavat mallin tai sen osan parametrien identifioitavuuden.

Hyvä lähtökohta ongelmaa tarkasteltaessa on pitää mielessä, että sekä a priori tieto (=mallimme) että a posteriori tieto (=havainnot) asettavat rajoituksia sille rakenteelle, jonka oletamme generoineen havainnot. Käyttämällä kummankinlaisia rajoituksia hyväksi pyrimme määräämään mahdollisimman pitkälle itse rakenteen. Mikäli malli ei identifioidu, emmekä tyydy ei-identifioituvaan malliin, joudumme yleensä tilanteeseen, jossa mallin identifioimiseksi siihen on lisättävä lisärajoituksia, ns. identifiointirajoituksia. Tämä merkitsee itse asiassa mallin muuttamista.

Eräs "leikkiprosessien" joukko, jonka ominaisuuksia on melko pitkälle tutkittu, ja jota tässä käytetään esimerkkinä edellä esitetylle yleiselle esitykselle, on ns. stokastiset lineaariset mallit. Rakennemuodossa malli voidaan kirjoittaa

$$A1. \quad A\underline{y}_t + B\underline{z}_t = \underline{u}_t \quad (2.1.)$$

$$A2 \quad K \quad (2.2.)$$

Tässä \underline{y} on riippuvien eli endogeenisten muuttujien vektori, \underline{z} ennaltamäärättyjen muuttujien vektori, \underline{u} virhetermien vektori, A on $n \times n$ riippuvien muuttujien kerroinmatriisi ja B $m \times n$ ennalta määrättyjen muuttujien kerroinmatriisi. Mukana on siis n riippuvaa ja m ennaltamäärättyä muuttujaa¹⁾.

K on joukko lauseita, joka sisältää kaiken muun kuin $A1:ssä$ ilmoitetun mallia koskevan a priori tiedon, esim. tiedon jonkin parametrin arvosta tai vaihteluvälistä, oletuksen $A:n$ singulaarisuudesta, siitä ettei ennaltamäärättyjen muuttujien välillä ole lineaarisia identiteettejä, siitä että en-

1) Yleensä muuttujat luokitellaan endogeenisiin ja exogeenisiin, mutta useissa tarkasteluissa (kuten nytkin) on tässä käytetty luokittelu käyttökelpoisempi. Riippuvia muuttujia ovat ainoastaan endogeenisten muuttujien arvot ko. ajankohtana (=hetkellä t), ennalta määrättyjä ovat eksogeeniset ja viivästetyt endogeeniset muuttujat. Mikäli mallissa ei ole mukana viivästettyjä endogeenisiä muuttujia mainitut kaksi luokittelua yhtyvät. Mikäli käyttäisimme ainoastaan jakoa endogeeniset eksogeeniset pitäisi malli esittää muodossa

$$\sum_{\tau=0}^p A_{\tau} y_{t-\tau} + Bx_t = \underline{u}_t$$

jossa \underline{y}_t on endogeenisten muuttujien vektori hetkellä t ja \underline{x}_t eksogeenisten muuttujien vektori. Tällöin eksogeenisten muuttujien viivästetyt arvot yksinkertaisuuden vuoksi määriteltä uusina eksogeenisinä muuttujina.

naltamäärätyt muuttujat ja häiriötermit ovat korreloimattomia, muut häiriötermien jakaumaa koskevat oletukset, ym. Tässä yhteydessä ei häiriötermejä koskeviin oletuksiin kuitenkaan tarkemmin puututa, koska identifiointiongelmia voidaan esitellä ilmankin. Huomautettakoon kuitenkin, että häiriötermeistä voidaan tehdä useita erilaisia oletuksia ja usein myös tällaisia rajoituksia voidaan käyttää tehokkaasti mallin tai sen osan identifioimiseen ((ks. liite II).

Lineaarisisessa mallissamme edellä esitettyä operaatiota I vastaava transformaatio on lineaarinen transformaatio. Jos F :n oletetaan olevan normaalisen, myös G on normaalin. Jos lisäksi $E(\underline{u}) = 0$ F :ää riittää karakterisoimaan kovarianssimatriisi Σ^u . Samoin G :tä voidaan karakterisoida $E(\underline{y})$:llä ja Σ^y :llä. Mallin ns. redusoitu muoto on $\underline{y} = -A^{-1}Bz + A^{-1}\underline{u}$ ja

$$\Sigma^y = A^{-1}\Sigma^u A^{-1} \quad (2.3.)$$

$$E(\underline{y}) = -A^{-1} Bz \quad (2.4.)$$

Tiettyä rakennetta S joka riippuu edellä kuvatulla tavalla häiriötermin jakautumasta ja operaatiosta I, voidaan nyt kuvata Σ^u :lla ja matriiseilla A ja B eli

$$S = (\Sigma^u, A, B) \quad (2.5.)$$

Edellä esitetty operaation I suorittaminen enteilee jo erästä lineaaristen mallien identifointi-ongelmaa: näemme että, mikäli Σ^u :sta ei ole vahvaa tietoa A^{-1} :n elementtien määräämiseen, voi olla että jälkikäteen emme pysty erottelamaan matriisista $\pi = A^{-1}B$ alkuperäisiä matriiseja A ja B . Onhan π :ssä vain $n \times n$ elementtiä verrattuna $n \times n + n \times m = n \times (n+m)$ elementtiin A :ssa ja B :ssä yhteensä.

2.2. A posteriori observoimalla saatava informaatio

Edellä esitetystä lienee jo käynyt selville, että mikäli emme ole valmiit etukäteen rajoittamaan millään lailla niiden selitysten joukkoa, jotka olemme valmiit hyväksymään, emme pysty saamaan mitään tietoa siitä mekanismista joka havaintoja tuottaa. Luonnollisesti saamme informaatiota, mutta sen arvo on rajoitettu, emmekä pysty järjestämään sitä laajempaan ilmiökenttään nähden.

Kun mallimme (sallittujen rakenteiden joukko) on lyöty kiinni on oleellista se, että havainnot asettavat rajoituksia mallillemme. Tämä on luonnollista, sillä väitämmehän juuri jonkun mallimme rakenteista generoivien havainnot. Otetaan esimerkiksi deterministinen lineaarinen malli

$$y = a + bx \quad (2.6.)$$

jossa y ja x havaittavia muuttujia ja a ja b ennalta määräämättömiä kiinteitä lukuja. Saatuaamme yhden havainnon (y_1, x_1) emme vielä voi määrätä lukuja a ja b , mutta olemme saaneet niiden välille riippuvuuden $a + x_1 b = y_1$ (tässä nyt y_1, x_1 kiinteitä lukuja ja a ja b tuntemattomia). Toinen havainto tuottaa ehdon $a + x_2 b = y_2$ ja tästä yhtälöryhmästä voimme määrätä tuntemattomat parametrit (oletetaan etteivät x_i ja y_i molemmat ole nolliä). Teoreettisesti apukäsitteet a ja b voidaan siis määrittellä havainnoittavien muuttujien x ja y avulla.

Huomattakoon, että mikäli edellä saamme kolmannen havainnon, saamme myös kolmannen yhtälön, jolloin kahden muuttujan yhtälöryhmässämme on kolme yhtälöä. Tietyt loogikot (Simon, ilmeisesti Blalock) kutsuvat tällaista tilannetta "yli-identifioituvaksi", koska meillä on enemmän kuin riittävästi havain-
toja mallin identifioimiseksi (ks. luku 4.).

Tarkasteltaessa yleisemmin lineaarisen mallin (2.1.)-(2.2.) suhdetta a posteriori tietoon voidaan todeta että havainnot deterministisessä tapauksessa asettavat m rajoitusta kunkin yhtälön parametreille¹⁾. Tämä johtuu havaintojen muodostamis-
tavasta: valitaan m:lle ennalta määrätyle muuttujalle arvot ja annetaan mallin määrätä riippuvien muuttujien arvot. Toisistaan yksityisiä yhtälöitä emme kuitenkaan pelkästään havaintojen perusteella voi erottaa. Yli m:stä havainnosta ei siis itse asiassa ole mitään hyötyä, mikäli emme epäile mallimme paikkansa pitävyyttä ja halua käyttää havaintoja sen testaamiseen. Edellä on luonnollisesti oletettava, että oletuksemme siitä että havainnot on todella generoitu mallillamme, on oltava voimassa.

Myös stokastisessa tapauksessa, tutkittaessa asiaa populaatio-
tasolla, asettavat havainnot ainoastaan m rajoitusta kunkin yhtälön parametreille. Lausejoukko K (mallin aksioomat A2.) sisältää usein oletuksen ennaltamäärättyjen muuttujien ja virhetermien korreloitumattomuudesta, ja tulos on johdettavissa tämän avulla.

1) Ks. esim. Fisher (6).

Erikoisesti huomattakoon, että vaikka lineaarisen mallin rakennemuoto ei olekaan identifioitavissa, on sen redusoitu muoto aina identifioituva, kun havaintoja on tarpeeksi. Redusoidun muodon matriisi on siis määrätty yksikäsitteisellä tavalla havaintojen perusteella. Mikäli emme ole kiinnostuneet rakennemuodosta, onkin eräisiin mallin käyttötarkoituksiin riittävää tuntea tämä. Edellä esitetty merkitsee, sitä että redusoidun muodon matriisiin π sisältyvät parametrit on eksplisiittisesti määriteltä havaintojen avulla.

2.3. A priori ja a posteriori tiedon yhdistäminen ja parametrien identifiointi

A posteriori informaatio kykeni asettamaan m rajoitusta kuhunkin yhtälöön liittyville parametreille, joita kaiken kaikkiaan (jos emme käytä hyväksi a priori informaatiota) oli yhtälössä $m+n$. Jotta kuhunkin yhtälöön tarvittavat parametrit voitaisiin määrätä, tarvitaan siis n lineaarista riippuvuutta lisää parametrien välille. Yksi tällainen ehto saadaan yleensä normalisointisäännöstä joka kertoo minkä muuttujan suhteen kukin yhtälö on normalisoitu, eli minkä muuttujan kerroin a_0 yhtälössä on 1¹⁾. Tätä normalisointia vastaa usein intuitiivinen käsityksemme siitä mitä muuttujaa kussakin yhtälössä useamman yhtälön simultaanimallin puitteissa selitetään. Lisäksi tarvitaan siis a priori rajoituksia $n-1$ kappaletta. (Muistettakoon että n oli endogeenisten muuttujien lukumäärä.)

1) Ks. Fisher (6).

Edellä olevalla yksinkertaisella päättelyllä olemme johtaneet tärkeän ns. "order-ehdon" joka lineaarisessa mallissa kertoo milloin tietty yhtälö (s.o. tietyn yhtälön parametrit) on identifioituva.

ORDER-EHTO: Lineaarisen mallin yhtälö on identifioituva, mikäli malli sisältää $n-1$ a priori lineaarista rajoitusta ko. yhtälön parametreille.

Edellä esitetty ehto on välttämätön, muttei riittävä ehto lineaarisen mallin (2.1.)-(2.2.) yhtälön parametrien identifioimiselle. Yksityiskohtaisemmilla tarkasteluilla on johdettavissa ns. rank-ehto, joka on myös riittävä¹⁾. Tässä ei syvennytä tilanteeseen tarkemmin; oleellista on ollut osoittaa, että useissa erilaisissa tilanteissa on mahdollista melko yleisesti samaa, milloin tietyn tyyppinen malli tai sen osa on identifioituva. Tällöin päädytään juuri yllä esitetynlaisiin sääntöihin.

Silloin kun lineaarinen malli ei identifioidu, on mahdollisia rakenteita yleensä ylinumeroituva määrä, joten n -ali-identifioituvat mallit ovat harvinaisia²⁾.

1) Rank-ehdosta ks. esim. Goldenberger (1), Christ (1).

2) Esimerkki 2-ali-identifioituvasta mallista löytyy esim. kirjasta Christ (1) ss. 337-338.

2.4. Rakennemuodon yhtälöiden lineaarikombinaatiosta

Lineaarisen mallin yhtälö on siis identifioituva, mikäli se rakennemuodon yhtälöistä on ainoa joka paitsi, että se on sopu-
soinnussa havaintojen kanssa, on myös sopu-
soinnussa niiden rajoitusten kanssa, jotka rakenteen kuulumisen tiettyyn malliin juuri tälle yhtälölle asettaa. Yksityisen yhtälön kertominen vakiolla muuttaa yhtälön parametrit, mutta oleellisesti yhtälö yleensä pysyy samana, esim. se on edelleen sopu-
soinnussa havaintojen kanssa¹⁾.

Yleisesti voidaan osoittaa intuitiivisesti melko selvältä tun-
tuva seikka, että "todellinen" yhtälö voidaan erottaa kilpaile-
vista "vääristä" yhtälöistä mikäli ko. rakennemuodon yhtälö ei
ole muiden rakennemuodon yhtälöiden lineaarikombinaatio²⁾. Tämä
pitää tietysti paikkansa kaikille rakennemuodon yhtälöille,
joten a posteriori informaatio ei siis voi erottaa todellisten
yhtälöiden joukkoa, eli todellista rakennetta, muista rakenteis-
ta jotka saadaan muodostamalla yhtälöt todellisten yhtälöiden
lineaarikombinaatioina. Olkoon E $n \times n$ nonsingulaarinen matriisi.
Kertomalla yhtälö (2.1.) vasemmalta E :llä saadaan

$$A^{\$}y_t + B^{\$}z_t = u_t^{\$} \quad (2.7.)$$

jossa $A^{\$} = EA$, $B^{\$} = EB$ ja $u_t^{\$} = Eu_t$.

1) Ks. kuitenkin Fisher (6) ss. 25-26.

2) Katso esim. Christ (1); Koopmans (1) tai Goldberger (1).

Pelkästään havainnoilla saatava informaatio ei siis pysty eroittamaan yhtälön (2.7.) liittyvää rakennetta yhtälöön (2.1.), liittyvästä rakenteesta koska molemmat rakenteet ovat sopusoinnussa havaintojen kanssa.

2.5. Identifioitavuus ja redusoitu muoto

Edellä tuli jo esille että huolimatta siitä ettei mallin rakennetta voida identifioida, on sen redusoidun muodon parametrit aina määrättävissä yksikäsitteisesti. Alkuperäisen yhtälösystemin rakennemuoto oli

$$A\underline{y}_t + Bz_t = \underline{u}_t \quad (2.8.)$$

ja vastaava redusoitu muoto

$$\underline{y}_t = \pi z_t + \underline{v}_t = (-A^{-1}B)z_t + A^{-1} \underline{u}_t \quad (2.9.)$$

Mikäli alkuperäinen malli muunnetaan edellisessä luvussa esitetyllä tavalla lineaarisella muunnoksella E, on uutta rakennetta vastaava redusoitu muoto

$$y_t = \pi^{\$} z_t + v_t^{\$} \quad (2.10.)$$

jossa kuitenkin

$$\pi^{\$} = -A^{\$-1}B^{\$} = -(EA)^{-1}(EB) = -A^{-1}B = \pi \quad (2.11.)$$

$$v_t^{\$} = A^{\$-1}u_t^{\$} = (EB)^{-1}(Fu_t) = A^{-1}u_t = v_t \quad (2.12.)$$

Kuten odottaa sopiikin on molempien rakenteiden redusoitu muoto siis sama. Tämähän on välttämätöntä, koska edellä oli osoitettu että alkuperäinen ja tästä lineaarimuunnoksella saatava rakenne ovat molemmat sopusoinnussa havaintojen kanssa, eli molempien tulee generoida samat $y:n$ havaintoarvot, kun $z:n$ ja $u:n$ arvot on annettu.

Koska mallin redusoitu muoto on aina identifioitavissa ja redusoidun muodon parametrit sisältävät yksinkertaisen lineaarisen mallin tapauksessa kaiken sen informaation mitä havainnoilla voidaan saada rakennemuodon parametreista - ne saadaan ei-singulaarisella transformaatiolla yhtälöistä jotka sisältävät kaiken tällaisen informaation - identifiointiongelma esitetään usein kysymyksenä voidaanko rakennemuodon parametrit määrätä yksikäsitteisesti, kun redusoidun muodon parametrit tunnetaan. Rakennemuodon parametreja vastaavat aina yksikäsitteiset redusoidun muodon parametrit.

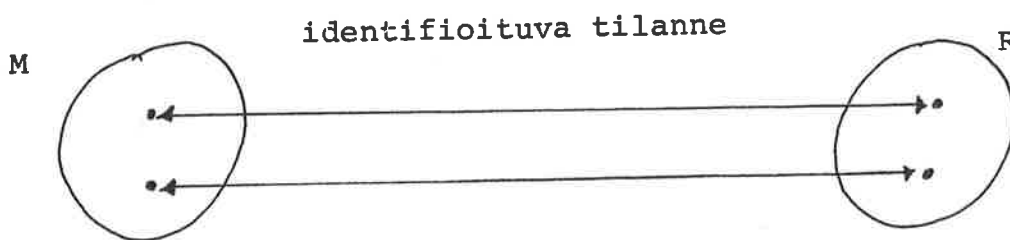
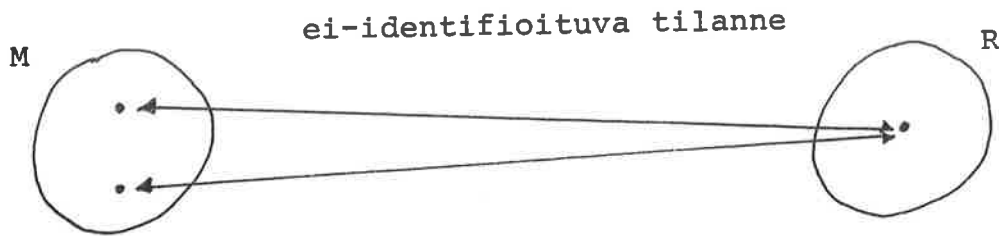
Ei-identifioituvassa tapauksessa redusoidun muodon parametreja eivät vastaa yksikäsitteiset rakennemuodon parametrit, vaan useammat erilaiset rakennemuodon parametrien kombinaatiot tulevat kysymykseen, emmekä osaa valita näiden välillä. Identifioituvassa tapauksessa riippuvuus on yksikäsitteinen molempiin suuntiin.

Malli

$$M = (\Sigma^u, A, B) \mid A \in \mathcal{A}, B \in \mathcal{B}, \Sigma^u \in \mathcal{Z}^u$$

Mallin redusoitu muoto

$$R = (\Sigma^v, \pi) \mid \pi = -A^{-1} B, \Sigma^v = A^{-1} \Sigma^u A^{-1}$$



Ns. "yli-identifioituvassa" tilanteessa ei stokastisissa lineaarisissa malleissa (todennäköisyydellä 1) johtuen äärellisestä otoksesta löydy lainkaan rakennetta, joka vastaisi estimoitua redusoitua muotoa. Tässä kohden kirjallisuudessa esiintyvä terminologia on hiukan epäjohdonmukainen ja hämäävä. "Yli-identifioituvat" mallit ovat aina identifioituvia. Kirjallisuudessa esiintyy kaksi erilaista "yli-identifioitavuuden" käsitettä, joita on käsitelty tuonnempana. Joka tapauksessa on muistettava, että "yli-identifioitavuus" on molemmissa tapauksissa aivan erilainen ilmiö kuin identifioitavuus. Olisikin paikallaan käyttää termin "yli-identifioituva" sijasta jotain muuta termiä, jolloin mallit ovat yksinkertaisesti joko identifioituvia tai eivät sitä ole.

3. ERÄÄSTÄ TERMIN "ALI-IDENTIFIOITAVUUS" KÄYTÖSTÄ

Malli, joka periaatteessa on identifioituva voi (mikäli hyväksymme eräiden kirjoittajien kielenkäytön) olla "ali-identifioituva", koska meillä ei ole riittävästi havaintoja eri parametrien määräämiseen. Esim. mallissa $y=a+bx$, jossa y ja x ovat havaittuja muuttujia ja a ja b tuntemattomia parametreja (teoreettisia apukäsitteitä) emme yhden havainnon perusteella pysty määräämään molempia parametreja, ellei meillä ole muuta a priori tietoa niistä, esim. tunnemme toisen parametrin arvon, tai niiden suhteen. (Tällaista terminologiaa käyttää mm. Simon).

Tuntuisi kuitenkin järkevämmältä myös tässä tapauksessa kutsua mallia periaatteessa identifioituvaksi, eli lyhyesti identifioituvaksi. Näin yleensä tehdäänkin ekonometrisessä kirjallisuudessa, esim. Marschak (1). Tämä käytäntö olisi myös paremmin sopu-soinnussa muun terminologian kanssa. Huomattakoon että edellä mainittu oletus äärettömästä otoksesta tutkittaessa identifioituvuutta stokastisissa malleissa merkitsee juuri sitä, että tutkitaan identifiointia periaatteellisessa mielessä. Oletetaan siis riittävä määrä havaintoja, eikä olla kiinnostuneita tämän tapaisesta "ali-identifioitavuudesta". Mikäli periaatteessa

identifioituva malli jäisi näin identifioitumatta, olisi tämä luonnollisesti erikseen selostettava, mikäli mahdollista käyttäen jotain muuta termiä kuin "ali-identifioitavuus".

Sinänsä edellä esitetynlainen teoreettisten käsitteiden avonaisuksi jääminen puutteellisten havaintojen vuoksi on erittäin mielenkiintoinen ilmiö ja kaipaisi mahdollisesti omaa termiä.

4. "YLI-IDENTIFIOITAVUUDESTA"

Tietyissä erikoistilanteissa identifioituvia malleja kutsutaan usein "yli-identifioituviksi". Termillä on kaksikin erilaista käyttöä. Molemmat vastaavat tilannetta, joka on oleellisesti erilainen ei-identifioitavuudesta, ja eri termien käyttö olisi paikallaan kummassakin tapauksessa. Ensimmäinen "yli-identifioitavuuden" käsite liittyy samantapaiseen tilanteeseen kuin luvussa 3 esitelty termi "ali-identifioituvuus". Mikäli meillä deterministisessä mallissa on enemmän kuin tarpeeksi havaintoja parametrien identifioimiseen, voimme määrätä parametrit useiden eri havaintoryhmien avulla, parametrit voidaan eräässä mielessä "yli-identifioida". Huomattakoon kuitenkin, että mikäli malli on sopusoinnussa empirian kanssa, tulevat parametrien arvot eri tavoilla laskettuna samoiksi. Itse määritelmähän on sama, ainoastaan eri havain-
toja käytetään yhdessä saman määritelmän kanssa arvioimaan apukäsitteiden arvoja.

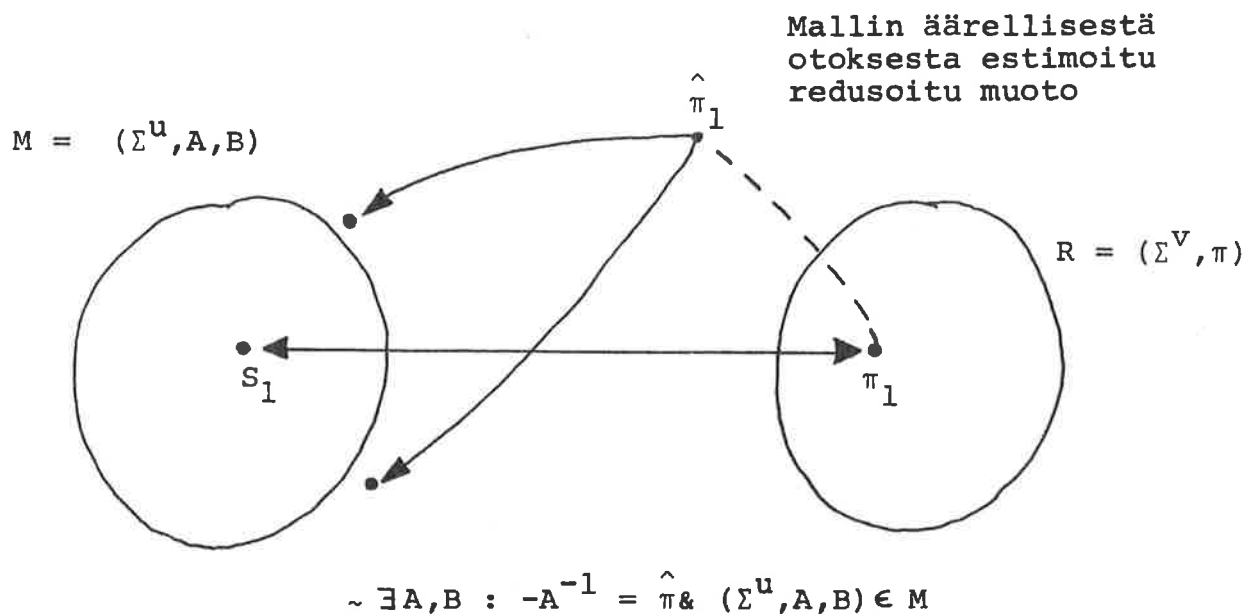
Edellä kuvatuslaista terminologiaa käyttää esim. Simon (kuten vastaavaa "ali-identifioitavuuden" käsitettäkin). Tämän tapaista "yli-identifioitavuutta" voidaan käyttää mallin testaamiseen, aivan samalla tavalla kuin yleensäkin testaamme

mallejamme havaintojen avulla. Niin kauan kuin mallissa riittää vapaita parametreja, ei mallia itse asiassa ole vielä testattu havaintojen avulla. Havaintoja on oltava vähintään niin paljon, että ainakin periaatteessa olisi mahdollista joutua ristiriitaiseen tilanteeseen.

Tällainen ristiriitainen tilanne merkitsee sitä, että mallimme on liian rajoittava, ei löydy yhtään rakennetta jolla havainnot voitaisiin ajatella generoidun. Mallimme on yksinkertaisesti väärä, ja se on hylättävä, ellemme tyydy likimääräiseen tarkkuuteen. Tällöin kuitenkin jää ongelmaksi mitä havaintoja käytetään parametrien määräämiseen, ja tilanne on melko lähellä (vain eräässä mielessä) ekonometrisessä kirjallisuudessa esiintyvää toista "yli-identifioitavuuden" käsitettä.

Tämä termin "yli-identifioitavuus" käyttö liittyy stokastisiin malleihin ja tilanteeseen, jossa äärellisen otoksen satunnaisvaihtelusta johtuen parametrit voidaan määrätä useammalla eri tavalla. Kyseessä on lähinnä estimointitekhninen ongelma ja ongelma häviää otoksen kasvaessa, esimerkiksi oletettaessa edellä esitettyyn tapaan endogeenisten muuttujien ehdollinen yhteisjakauma tunnetuksi. Sen apparaatin puitteissa, jota luvussa 1. käytettiin identifioitavuuden määrittelyyn, "yli-identifioitavuutta" tässä mielessä on vaikea käsitellä. Huomautettakoon kuitenkin, että kaikki "yli-identifioituvat" mallit ovat siellä annetun määritelmän mukaan identifioituvia malleja.

Hyvä esimerkki "yli-identifioitavuudesta" on tilanne, jossa lineaarisen mallin äärellisestä otoksesta estimoitu redusoitu muoto asettaa liikaa rajoituksia mallin "rakennekandidaateille" ja yhtään sellaista rakennetta, johon liittyvät matriisit A ja B johtaisivat estimoituun redusoidun muodon matriisiin $\hat{\pi}$ ei löydy mallista M. Luonnollisesti tällaisia rakenteita löytyy mallimme ulkopuolelta. Tilannetta on tarkasteltu oheisessa kuviossa.



Korostettakoon vielä että vastaavuutta ei nimenomaan löydy S_1 :n ja $\hat{\pi}_1$:n välillä. Kun otoskoko kasvaa, ja mikäli havainnot on todella generoitu jollakin identifioituvan mallimme rakenteista, esim. S_1 , löytyy lopulta kyllä yksikäsitteinen vastaavuus rakennemuodon ja "oikean" redusoidun muodon π_1 välillä. "Yli-identifioitavuutta" voidaan verrata tilanteeseen, jossa arvioimme suomalaisten 21-vuotiaiden miesten keskipituutta käyttäen useampia otoksia, joissa otoskeskiarvot eroavat. Myös saman otoksen puitteissa voidaan käyttää samalle teoreettiselle

populaatioparametrille erilaisia (konsistentteja) estimaattoreita. Tällainen on tilanne, jos esimerkiksi oletamme miesten pituuksien jakaantuneen normaalisesti ja käytämme äärellisessä otoksessa aritmeettista keskiarvoa ja mediaania populaatiokeskiarvon arvioimiseen. Mikäli otos on ääretön ja havainnot on todella poimittu normaalijakaumasta, häviää ongelmamme, samoin kuin "yli-identifioitavuus" lineaarisissa stokastisissa malleissa. Eräs vaikeus käytännössä on tietenkin se että ennemmin tai myöhemmin leikkiprosessimme ei riitä kuvaamaan todellisuutta ja näin ko. tuloksella ei olisi käyttöä, vaikka saisimme äärettömiäkin otoksia. Se, että hallinnassamme on vain eräs leikkiprosessi paljastuisi näet armotta.

Silloin kun käsitellään äärellisiä otoksia ja halutaan jakaa identifioituvat mallit "yli-identifioituviin" ja muihin malleihin, kutsutaan näitä malleja usein "täsmälleen" identifioituviksi. Huomattakoon, että koko lineaaristen mallien estimointiteoria voidaan jaoitella sen mukaan käsitelläänkö ei-"täsmälleen" -, vai "yli-identifioituvia" malleja. Ei-identifioituvien mallien rakenneparametreja ei voida lainkaan estimoida (mutta monia näiden parametrien funktioita voidaan, esim. redusoitu muoto). "Täsmälleen" identifioituvan mallin kaikki parametrit ovat yksinkertaisella tavalla estimoitavissa (esim. ns. indirect least squares). Suuri osa lineaaristen mallien estimointiteoriaa käsittelee juuri "yli-identifioituvaa" tapausta, jossa on vaikeutena päästä yksikäsitteisiin rakennemuodon parametreihin siten että kaikki a priori tieto tulisi efektiivisesti käytettyä ja siten, että kaikki a priori rajoitukset olisivat voimassa.

5. HUOMIOITA IDENTIFIOITAVUUDESTA JA MÄÄRITELTÄVYYDESTÄ

Kun sanotaan: "yli-identifioituvassa" tilanteessa voidaan teoreettiset apukäsitteet määritellä usealla tavalla" ja kun sanotaan "n-ali-identifioituvassa tilanteessa voidaan teoreettiset apukäsitteet määritellä usealla tavalla" on todella kyse eri "tavoista".

"Yli-identifioituvassa" tilanteessa meillä on useita tapoja määritellä sama käsite samassa mielessä kuin 4 voidaan määritellä $+2$ tai $-2:n$ toisena potenssina tai populaatiotasolla keskiarvo ja mediaani kaikissa mahdollisissa normaalijakautumissa johtavat samaan lukuun. Tavat ovat ekvivalentteja siinä mielessä, että sama "maailman rakenne" saadaan kiinni usealla tavalla.

n-ali-identifioituvassa tapauksessa meillä on samalle teoreettiselle apukäsitteelle avoinna useita keskenään ristiriitaisia määritelmiä, joista kuitenkin voimme väittää jonkin olevan voimassa. Tällainen on tilanne jos esim. kahden parametrin summa on määrätty, mutta itse parametrien arvot voivat ao. rajoituksen puitteissa vaihdella, siis mahdollisia

"maailman rakenteita" on useampia. Tässä mielessä n-ali-identifioitavuus ja yleensä ei-identifioitavuus vastaa eksplisiittisten määritelmien disjunktiota (so. jokin annetuista määritelmistä pitää "todelliselle rakenteelle" paikkansa). Tällöin useammat rakenteet, (jotka kuitenkin kaikki ovat observationaalisesti ekvivalentteja) ja näin useammat teoreettisten apukäsitteiden eli parametrien arvot ovat mahdollisia.

Yli-identifioitavuus tulee nähdäkseni vastaamaan eksplisiittisten määritelmien konjunktiota. "Todellisessa rakenteessa" pitävät kaikki yli-identifioituvan parametrin eksplisiittiset määritelmät paikkansa. Yli-identifioitavuus vastaa näin ollen mielenkiintoista, erittäin voimakasta määriteltävyyden lajia.

"Yli-identifioitavuus" (Simonin mielessä) esim. ylimääräisten havaintojen kautta vastaa eksplisiittistä määriteltävyyttä, mikäli malli periaatteessakaan on sopusoinnussa havaintojen kanssa. Mikäli malli taas ei sovi kuvaamaan todellisuutta eli eri havainnoista saadaan apukäsitteille ristiriitaiset arvot (määritelmistä tuskin sinänsä voidaan käyttää sanaa ristiriitaiset) on malli yksinkertaisesti väärä, eikä apukäsitteiden määriteltävyys enää ole niin mielenkiintoista.

Mikäli taas on kyse "yli-identifioitavuudesta" joka esim. stokastisissa malleissa johtuu pienestä otoksesta ja siitä, että meillä on useita menetelmiä joilla teoreettiset apukäsitteet (sinänsä konsistentisti, mutta äärellisen otoksen satunnaisvaihtelusta johtuen ristiriitaisilta näyttävillä

tavoilla) voidaan määrätä useammalla tavalla, ollaan logiikan tavanomaisen, lähinnä deterministisiin tapauksiin soveltuvan määriteltävyyden käsitteen ulkopuolella. Populaatiotasollahan ristiriidat häviävät ja huomataan että itse asiassa meidän määritelmämme ovat tällä tasolla operoivan tarkastelukehikon puitteissa ekvivalentteja.

Runsaasti mielenkiintoisia määriteltävyyden ongelmia olisi varmasti löydettävissä tarkastelemalla ekonometrian tutkimusmenetelmiä tästä näkökulmasta. Esimerkiksi ehdolliselle määriteltävyyden käsitteelle olisi luultavasti löydettävissä useitakin vastineita. Eräät ilmiöt ovat usein havaittavissa vain tietyissä olosuhteissa, esim. nousu- tai laskusuhdanteen aikana jne. Dummymuuttujien käyttö, samoin kuin tietynlaiset kausipuhdistusmenetelmät, lienevät myös hyviä esimerkkejä. Mielenkiintoinen on myös regressioanalyysissä tilanne, jossa yhtä muuttujaa selitetään vaikkapa havaintojen puutteellisuuden johdosta muiden muuttujien pääkomponenteilla (vrt. Tuomela (1) ja faktorianalyysi). Monimuuttujamenetelmät tarjoavat samantapaisia ongelmia useita.

Taloudellisessa kirjallisuudessa esiintyy myös täysin verbaalisella tasolla apukäsitteitä, joiden määrittäminen on ongelmallista. Tällaisia ovat esim. investointiteoriassa tärkeä "odotetut tuotot". Tämä tulevaisuuden odotuksia kuvaava apukäsite joudutaan kuitenkin määräämään jollain tavalla historiallisen hinta-, kustannus-, tuotanto- ym. kehityksen avulla, eli määrittämään eksplisiittisesti suoraan observaatiotermien avulla.

6. HUOMIOITA TEOREETTISTEN APUKÄSITTEIDEN TARPEELLISUUDESTA

Seuraavien esimerkkien ja niiden yhteydessä esitettyjen kommenttien tarkoitus on lyhyesti sanottuna perustella seuraavaa väitettä: Rakennemuodon parametrien (teoreettisten apukäsitteiden) ja niiden identifioituvuuden (eksplisiittisen määriteltävyyden) tarve riippuu oleellisesti siitä tarkoituksesta mihin ko. mallia (teoriaa) aiotaan käyttää.

Tarkastelkaamme ensimmäisenä apukäsitteiden ja niiden identifioituvuuden tarvetta koskevana esimerkkinä Elinkeinoelämän Tutkimuslaitoksessa rakennetun mallin käyttöä vuoden 1973 alussa eri tutkimuslaitoksissa suoritetuissa, tulopolitiikkaa koskevissa arvioissa.

Malli on estimoitu rakennemuodossa, joten sen kaikki rakennemuodon parametrit tunnetaan, ja näiden perusteella voidaan aina laskea mallin redusoitu muoto π_0 jota käytetään ennustamiseen. Erikoisesti jotkut mallin yhtälöt kuvaavat taloudellisten agenttien toimintaa, kulutusta, investointia, hinnanmuodostusta, ansiotasonmuodostumista, ulkomaankauppaa jne.

Näin esim. ansiotason määräytymiseen liittyy tietty mekanismi, joka on määrätty estimointiperiodilla havaitun käyttäytymisen avulla. Yhtälö kertoo, minkälainen vaikutus hintakehityksellä, tuottavuudella ja työttömyydellä on ansiotasoon.

Mikäli oletamme että ko. mekanismi ja vastaavasti mallin muu rakenne myös ennustetilanteessa on voimassa, voimme käyttää ennustetarkoituksiin matriisia π_0 . Jos nyt kuitenkin ansiotason määräytymismekanismi oleellisesti muuttuu esim. siten (kuten keväällä 1973 mm. ehdotettiin) että tulopolitiikan keinoin voidaan määrätä ansiotason nousu reaalisesti tuottavuuden nousua vastaavaksi, on rakenteemme muuttunut ja joudumme ennuste- ja analyysitarkoituksiin käyttämään toista redusoidun muodon matriisia π_1 . Tämän voimme kyllä helposti määrätä, koska tunnemme uuden rakennemuodon. Mikäli alunperin emme olisi määränneet alkuperäistä rakennetta vaan olisimme pyrkineet ratkaisemaan ongelman suoraan redusoidun muodon π_0 avulla, emme pystyisi mallillamme tarkastelemaan tämän tapaisia ongelmia.

Edellä kuvatuunlaisia ongelmia, joissa uuden tiedon lisääminen malliin käy paremmin, mikäli tunnemme rakennemuodon, on helppo kuvitella runsaasti.

Huomattakoon, että redusoidun muodon käyttäminen merkitsee eräiden teoreettisten apukäsitteiden korvaamista lähempänä observaatiotermejä (periaatteessa aina näiden avulla eksplisiittisesti määriteltävissä olevilla) olevilla käsitteillä.

Keskustelu teoreettisten apukäsitteiden tarpeellisuudesta ja niiden määriteltävyydestä palautuu siis osin vanhaan keskusteluun siitä tarvitaanko rakennemuodon parametrien tuntemusta, ja jos, niin millaisissa tilanteissa. Mikäli teoreettisten apukäsitteiden arvot edellä oletettiin kiinteiksi, ei niitä aina ollut välttämätöntä määritellä eksplisiittisesti, mikäli käytämme teoriaamme puhtaasti ennustetarkoituksiin.

Esimerkkinä ei-identifioituvien mallien käytöstä mainittakoon tuttu tarina kahdesta vanginvartijasta, joista toinen puhui totta, toinen valehteli. Vangitulle esitetään kaksi ovea, joista toinen vie vapautteen ja toinen kuolemaan, ja hänen annetaan tehdä yksi kysymys vartijoille, johon nämä voivat vastata vain myöntävästi tai kieltävästi. Vangittu ei kuitenkaan tiedä kumpi valehtelee, kumpi puhuu totta.

Juttu voidaan formalisoida yksinkertaisesti seuraavalla tavalla:

$$y = abx \qquad (6.1.)$$

jossa a ja b ovat vanginvartijoiden sisäisiä tiloja kuvaavia parametreja (saavat arvon -1 tai $+1$ sen mukaan valehtelee ko. vartija vai ei) x on asiain todellista tilannetta kuvaava muuttuja ja saa arvot v tai $-v$ sen mukaan viekö ovi vapautteen vai ei. y on vanginvartijan vastausta kuvaava muuttuja, joka siis myös voi saada arvoja v tai $-v$.

Hyvä kysymys: Vastaisiko tuo toinen vartija myöntävästi, mikäli kysyisin häneltä viekö vasemmanpuoleinen ovi vapauteen? (Tällöin on oletettu, että molemmat vartijat tuntevat tilanteen täysin). Edellä esitetty malli on hyvä ennustetarkoitukseen, muttei rakenteen analysointiin. Teoreettisia parametreja a ja b ei tarvitse, eikä voida identifioida, mikäli havaintojen tekotapa on edellä kuvatulla tavalla rajoitettu. Käytännössä tällaisia rajoituksia usein liittyy havaintojen tekoon.

Parametreja a ja b ei tarvitse identifioida, koska olemme a priori voineet identifioida erään niiden funktion, tiedämme nimittäin, että $ab = -1$. Käyttötarkoitukseen (vapauteen pääseminen) riittää mallin yleinen tunteminen ja sen erään redusoidun muodon tunteminen.

Voidaan tietysti väittää, että itse asiassa tällöin käytämme mallia

$$y = -x \qquad (6.2.)$$

mutta ko. tilanteen pohjalta on paljon helpompaa ja luonnollisempaa konstruoida malli (6.1.).

Toinen yksinkertainen esimerkki mallista jossa teoreettisten apukäsitteiden eksplisiittinen määrittely saattaa tuottaa vaikeuksia on tyyppiä

$$y = ax + bx \quad (6.3.)$$

olevat mallit (x ja y havaittavia muuttujia, a ja b teoreettisia apukäsitteitä). Tuntuu ilmeisesti kuitenkin järkevältä käyttää tällaista mallia, eikä mallia

$$y = cx \quad (6.4.)$$

mikäli x :n vaikutus y :hyn on intuitiivisesti selvällä tavalla jaettavissa kahteen komponenttiin.

Eräs vastaavanlainen tilanne, jossa mieluummin jaamme yksinkertaisen, mutta epälineaarisen riippuvuuden kahteen komponenttiin on arvo, joka muodostuu määrä kertaa yksikköhinnasta vähennettynä alennuksella. Syynä lienee se, että ajattelumme pyrkii lineaarisuuteen. On helpompaa ja tutumpaa muodostaa ensin arvo määrän lineaarisena funktiona ja päätellä sen jälkeen alennus erikseen.

$$\text{Arvo} = \text{hinta} \times \text{määrä} - \text{alennus}$$

$$A = pxq - f(q)$$

Edellä on merkitty alennus määrän funktioksi (tukkualennus). Todellisuudessa lienee alennus myös hinnan funktio, sillä myyjät usein nostavat ensin hintoja ja antavat sitten suuren alennuksen. He myyvät enemmän käyttäen hyväksi ihmisten yksinkertaista käsitystä tilanteesta. Joka tapauksessa systeemi on toivottoman ei-identifioituva, mikäli voidaan havainnoida vain arvoja ja määriä.

Mainittakoon tässä yhteydessä useissa fysikaalisissa ilmiöissä käytetyt mallit, joissa varsinainen malli pitää paikkansa "ideaali-ilmiölle" ja sovellettaessa mallia käytäntöön arvioidaan poikkeamat tästä tietyillä korjauskertoimilla. Tuloksena on yleensä teorioita, jotka suurin piirtein pitävät paikkansa koko alueella (tai kaikissa tapauksissa) ja tarkennettua versiota käytetään tarpeen tullen. Tällöin saavutetaan lisäksi tietty ajattelun yksinkertaisuus, koska ei tarvita useampia perusteellisesti erilaisia malleja.

Edellä oli jo puhetta siitä, että ekonometriset mallit ovat eräänlaisia leikkiprosesseja tuntemattomien reaali-prosessien kuvaamiseen. On lisäksi ilmeistä, että pyrkiminen identifioituihin malleihin vaatii leikkiprosessilta riittävää yksinkertaisuutta. Ensinnäkään leikkiprosessissa emme voi ottaa mukaan kaikkia käsitettävien muuttujien välisiä riippuvuuksia (tällöin esim. luvussa 2.3. esitetty order-ehto ei laajemman systeemin puitteissa ehkä ole voimassa). Toisaalta riippuvuussuhteiden identifioiminen vaatii sen yksinkertaistamista esim. osa selittävästä muuttujista on suljettava pois. Jos siis hyväksymme väitteen, että reaali-prosessit ovat ei-identifioituvia, näyttää siltä, että pyrkiminen määriteltävissä olevien teoreettisten apukäsitteiden käyttämiin ainakin joissakin tapauksissa merkitsee, jos ei nyt suoranaista väärinspesifiointia, niin ainakin karkeammin todellisuutta approksimoiviin leikkiprosesseihin.

Argumentin tekee teoreettisten apukäsitteiden eksplisiittisen määriteltävyyden vaatijalle kiusalliseksi se, että vaikkei rakennemuoto identifioidukaan, identifioituu redusoitu muoto aina, eikä ei-identifioituvien mallien käyttö näin edellytä ko. approksimointia.

Voidaan kuitenkin osoittaa, että mikäli hyväksymme sen, että leikkiprosessit approksimoivat reaaliprosesseja "the proper question is not whether certain parameters assumed to be zero are in fact so, the issue is rather whether they are in some sense sufficiently small - whether the restriction that the corresponding variables do not appear is or is not a good approximation". Fisher (2).

Samoin laajempaa yhtälösystemiä osiin hajoitettaessa suoritettava muuttujien eksogenisoinen (ja mallin identifioiminen tämän avulla) on perusteltavissa, mikäli ne reaaliprosesseissa ovat tämän osajärjestelmän kannalta approksimatiivisesti eksogeenisiä.

The use of a priori restrictions which are approximate thus leads to negligible inconsistencies in such estimates also - for good enough approximations. It follows that the restricted estimates of the reduced form obtained from structural equation estimates converge more rapidly to probability limits that differ slightly or negligibly from the true reduced form coefficients than the unrestricted least squares reduced form estimates converge to the true reduced form parameters. Fisher (2).

Se minkälaisen kannan otamme apukäsitteiden tarpeeseen ja minkälaisia vaatimuksia asetamme niiden määriteltävyydelle riippuu siis oleellisesti siitä kuinka "tosikkomaisesti" suhtaudumme mallimme kykyyn kuvata empiriaa.

7. LOPPUPÄÄTELMIÄ

On ilmeistä, että määriteltävyyden ongelmat liittyvät läheisellä tavalla identifioitavuuden käsitteeseen. Koska asioita voidaan tarkastella erilaisten kehikoiden puitteissa ja kiinnittää huomiota erilaisiin seikkoihin, ei täsmällistä vastaavuutta eri kehikoiden käsitteiden välille ole toistaiseksi löydetty. On myös varauduttava siihen, ettei tällaista vastaavuutta joka kohdassa ole löydettävissäkään.

Identifioitavuus on vain eräs useista samantapaisia kysymyksiä käsittelevistä, eri tieteiden parissa ja eri yhteyksissä esiintulevista ongelmista. Tuntuisi luonnolliselta, että em. ongelmat pyrittäisiin palauttamaan logiikan puolella suoritettaviin määriteltävyyteen liittyviin tarkasteluihin.

Pyrittäessä palauttamaan identifioitavuutta määriteltävyyteen on erikoisena ongelmana se, että useat mielenkiintoiset seikat jäävät populaatiotasolla tarkastelun ulkopuolelle. Tähän saakka käytettyjen kehikkojen puitteissa voivat logiikan puolella suoritettut tarkastelut muodostaa näin ollen vain pohjan ekonometrisille identifiointitarkasteluille.

Korostettakoon lopuksi, että

- joissain tapauksissa emme kertakaikkiaan voi ratkaista ongelmiamme rajoittumalla malleihin, joissa apukäsitteet ovat eksplisiittisesti määriteltävissä. Ajattelumme yksinkertaisesti joskus toimii käyttäen ei-identifioituvia malleja. Osasyyn tähän on ilmeisesti se, että malliemme laajentuessa identifioituvat mallit muuttuvat ei-identifioituviksi. Jatkossa kuitenkin identifioimme tällaiset mallit käyttämällä operationaalisia vastineita käsitteillemme, jotka vähitellen alkavat elää hyvinkin itsenäistä elämää.
- toisaalta tietyn tyyppiset probleemat vaativat aina apukäsitteisen eksplisiittistä määriteltävyyttä.

Pohdittaessa edellä esitettyjä kysymyksiä tuntuisi esimerkiksi loogikkojen, tilastotieteilijöiden ja kansantaloustieteilijöiden yhteistyö luonnolliselta. Tällaisia hedelmällisiä keskusteluja voitaisiin ilmeisesti melko pienellä vaivalla Suomessa järjestää.

KIRJALLISUUTTA:

Christ, C.F. (1): *Econometric Models and Methods*, John Wiley and Sons, Inc., New York 1966, 705 s.

Fisher, F.M. (2): On the Cost of Approximate Specification in Simultaneous Equation Estimation, *Econometrica*, vol, 29, No. 2, April, 1961.

Fisher, F.M. (3): Identifiability Criteria in Nonlinear Systems: *Econometrica* vol. 29, No. 4, Marraskuu 1961 ss. 574-590.

Fisher, F.M. (4): Identifiability in Nonlinear Systems: A Further Note, *Econometrica*, Vol. 33, No. 1, Tammikuu 1965, ss. 197-205.

Fisher, F.M. (5): Near-Identifiability and the Variances of the Disturbance Terms, *Econometrica*, Vol. 33, No. 2, Huhtikuu 1965, ss. 409-419.

Fisher, F.M. (6): *The Identification Problem in Econometrics*, Mc Graw-Hill Book Company, New York 1966, 203 s.

Goldberger (1): *Econometric Theory*, John Wiley & Sons, New York, 1964.

Hurwicz, Leonid (1): Generalization of the Concept of Identification, *Statistical Inference in Dynamic Economic Models*, Cowles Commission, John Wiley & Sons, Inc., New York, 1950.

Hurwicz, Leonid (2): On the Structural Form of Interdependent Systems, *Symposium on Models in Empirical Sciences*.

Koopmans (1): Identification problems in Econometric Model Construction, Chapter 2 in *Studies in Econometric Method*, Cowles Commission Monograph 14, W.C. Hood and T.C. Koopmans, editors, New York: John Wiley and Sons, 1953, pp 27-48.

Koopmans, T.C., Rubin, H. and Leipnik, R.B. (2): Measuring the equation systems of dynamic economics, Chapter 2 in Statistical Inference in Dynamic Economic Models, Cowles Commission Monograph 10, editor T.C. Koopmans, New York, John Wiley and Sons, 1953 pp. 112-199.

Marschak, J. (1): Statistical Inference in Econometrics: An Introduction, Cowles Commission Monograph No. 10, John Wiley & Sons, pp. 1-45, 1953.

Liu, Ta-Chung (1): Underidentification, Structural Estimation and Forecasting, *Econometrica*, Vol. 28, No. 4, October, 1960.

Przelecki, M. (1): The Logic of Empirecal Theories, Routledge & Paul, London 1969.

Restle & Greeno (1): Introduction to Mathematical Psychology, Addison-Wesley Publishing Company, Reading, Massachusetts, 1970.

Simon, H. (1): Definable Terms and Primitives in Axion Systems, Symposium on the Axiomatic Method ss. 443-453.

Simon, H. (2): The Axiomatization of Physical Theories, Ilmeisesti julkaisematon moniste, saatu prof. Tuomelalta, Helsingin yliopisto, Käytännöllisen filosofian laitos.

Tuomela, R. (1): Identifiability and Definability of Theoretical Concepts, *Ajatus* XXX, 1968.