

# Keskusteluaiheita Discussion papers

Yrjö O. Vartia:

GAMMAJAKAUMAN  
PARAMETRIEN MAXIMUM-  
LIKELIHOOD ESTIMAATTOREIDEN  
APPROKSIMOINNISTA

No 6

10.10.1976

This series consists of papers with limited circulation, intended to stimulate discussion. The papers must not be referred or quoted without the authors' permission



Gammajakauman parametrien maximum-likelihood estimaattoreiden  
approksimoinnista

1. Johdanto

Kaksiparametrinen gammajakauma on normaali- ja betajakauman tapaan eräs matemaattisen tilastotieteen tärkeimmistä jatkuvista jakaumista. Gammajakaumalle käytän symbolia  $G(\lambda, \mu)$ , jossa  $\lambda$  ja  $\mu$  ovat positiivisia parametreja. Gammajakauman tiheysfunktio on

$$(1) \quad G(x|\lambda, \mu) = \frac{\lambda^\mu}{\Gamma(\mu)} x^{\mu-1} e^{-\lambda x} \varnothing(x>0), \text{ jossa}$$

$$(2) \quad \Gamma(\mu) = \int_0^\infty x^{\mu-1} e^{-x} dx, \text{ Eulerin gammafunktio,}$$

ja  $\varnothing(A)$  on ehdon  $A$  indikaattorifunktio ts.  $\varnothing(A) = 1$ , jos  $A$  on tosi ja  $\varnothing(A) = 0$ , jos  $A$  on epätosi. Merkintä  $\underline{x} \sim G(\lambda, \mu)$  tarkoittaa, että satunnaismuuttuja  $\underline{x}$  noudattaa gammajakaumaa parametrein  $(\lambda, \mu)$ . Parametri  $\lambda$  on skaalaustekijä ja  $\mu$  on käänteisessä suhteessa gammajakauman suhteelliseen hajontaan, esim. jos  $\underline{x} \sim G(\lambda, \mu)$  niin  $c\underline{x} \sim G(c\lambda, \mu)$ , kun  $c>0$ .

Gammajakauman  $r$ :s origomomentti  $E(\underline{x}^r)$  on

$$(3) \quad \Gamma(\mu+r)/\lambda^r \Gamma(\mu),$$

joten odotusarvoksi ja varianssiksi saadaan

$$(4) \quad E\underline{x} = \mu/\lambda,$$

$$(5) \quad D^2(\underline{x}) = E(\underline{x}^2) - (E\underline{x})^2 = \mu/\lambda^2.$$

Variaatiokerroin  $Dx/Ex = 1/\sqrt{\mu}$  riippuu ainoastaan parametrista  $\mu$  ja pienenee  $\mu$ :n kasvaessa. Geometrisen keskiarvon esitys

$$(6) \quad f = e^{\psi(\mu)}/\lambda$$

saadaan yhtälöistä

$$(7) \quad \log f = E \log x = \psi(\mu) - \log \lambda.$$

Tämä jälkimmäinen osa seuraa gammafunktion esityksestä

$$\Gamma(\mu) = \lambda^\mu \int_0^\infty x^{\mu-1} e^{-\lambda x} dx \quad \text{derivoimalla } \mu\text{:n suhteen: } D_\mu \Gamma(\mu) = \log \lambda \cdot \Gamma(\mu) + \Gamma(\mu) E \log x.$$

Gammajakaumalla on seuraava esim. Raon (1968, s. 133) todistama yhteenlaskuominaisuus: Olkoon  $x_1, \dots, x_k$  toisistaan riippumattomia satunnaismuuttujia ja  $x_i \sim G(\lambda, \mu), i = 1, \dots, k$ . Tällöin  $(x_1 + \dots + x_k) \sim G(\lambda, \mu_1 + \dots + \mu_k)$ . Tulos ei päde ellei parametri  $\lambda$  ole jokaiselle  $x_i$ :lle sama.

Jos  $\lambda = 1/2$  ja  $\mu = n/2$  saadaan gammajakauman erikoistapauksena  $\chi^2$ -jakauma vapausasteella  $n$ :  $\chi^2(n) = G(1/2, n/2)$ . Usein ajatellaan  $\chi^2$ -jakauma määritellyksi vain  $n$ :n kokonaislukuarvoilla, jolloin  $\chi^2(n)$  on  $n$ :n riippumattoman ja normaalijakauman  $N(0,1)$  noudattavan muuttujan  $z_i$  neliöiden summan jakauma:  $(z_1^2 + \dots + z_n^2) \sim \chi^2(n)$ .

Asettamalla  $\mu = 1$  saadaan gammajakauman erikoistapauksena eksponenttijakauma,  $\text{Exp}(\lambda) = G(\lambda, 1)$ , jota käytetään usein odotusajan jakaumamallina jonoteoriassa. Jos  $x_i$  on tietynlaisen tapahtuman odotusaika ja eri  $x_i$ :t ovat toisistaan riippumattomia sekä  $x_i \sim \text{Exp}(\lambda)$ , niin  $k$ :n tapahtuman odotusajan  $x_1 + \dots + x_k$  jakauma on yhteenlaskuominaisuuden nojalla  $G(\lambda, k)$ .

Gamma- ja betajakaumat ovat myös lähisukulaisia: Jos  $x$  ja  $y$  ovat toisistaan riippumattomia satunnaismuuttujia, joille  $x \sim G(\lambda, \mu)$  ja  $y \sim G(\lambda, \mu_2)$ , niin  $x/(x+y) \sim \text{Beta}(\mu_1, \mu_2)$ .

Esim. Cramér (1946, s. 371) on järjestettyjen otoksien tarkasteluissa soveltanut seuraavaa tulosta:  
Jos  $z \sim \text{Beta}(a, b)$ , niin muuttujan  $x=bz$  jakauma lähestyy gammajakaumaa  $G(a, 1)$ , kun  $b$  kasvaa rajatta. Tämän perusteella hyvin vinojen betajakaumien ominaisuuksia voidaan johtaa gammajakaumien avulla, mitä olen soveltanut julkaisussa Vartia (1973, s. 79-85).

Suuren taipuisuutensa vuoksi gammajakaumaa voidaan käyttää hyvin erilaisten positiivisten muuttujien jakaumamalleina. Jonoteoriaan liittyvien sovellutusten lisäksi sitä on ehdotettu viime aikoina esim. tulojen jakaumamalliksi, Salem & Mount (1974).

Gammajakauman soveltamista kuitenkin vaikeuttavat parametrien estimoinnissa vastaan tulevat epälineaariset yhtälöryhmät, joiden ratkaisemisessa on turvauduttu iteratiivisiin menetelmiin tai ratkaisujen taulukointiin, esim. Choi & Wette (1969). Vastaavia ongelmia syntyy laskettaessa eräitä harvinaisempia tunnuslukuja, ks. McDonald & Jensen (1975). Seuraavassa tarkastellaan gammajakauman maximum-likelihood yhtälöiden likimääräistä ratkaisemista käyttäen suhteellisen yksinkertaisia otossuureiden funktioita. Vastaavaa ongelmaa olen aiemmin tarkastellut myös betajakauman tapauksessa, Vartia (1973).

## 2. Gammajakauman parametrien maximum-likelihood yhtälöt

Olkoon  $(x_1, \dots, x_n)$  riippumaton otos jakaumasta  $G(\lambda, \mu)$ . Otoksen likelihoodfunktion avulla saadaan helposti likelihood-yhtälöt (ks. Choi & Wette (1969))

$$(7) \quad \psi(\mu) - \log \lambda = \log G$$

$$(8) \quad \mu/\lambda = \bar{x},$$

jossa  $\psi(x) = d \log \Gamma(x) / dx$  on psii- eli digammafunktio (ks. Abramowitz & Stegun (1970) tai Vartia (1973)),  $G = (x_1 x_2 \dots x_n)^{1/n}$  on havaintojen geometrinen keskiarvo ja  $\bar{x} = (x_1 + \dots + x_n)/n$  aritmeettinen keskiarvo.

Havaitaan, että yhtälö (7) merkitsee gammajakauman geometrisen keskiarvon  $f$  logaritmin (6) ja otoksesta määrätyn geometrisen keskiarvon logaritmin  $\log G$  yhtä suuriksi, ts. ml-menetelmä vaatii, että  $f = G$ . Vastaavasti (8) merkitsee odotusarvon (4) ja aritmeettisen keskiarvon  $\bar{x}$  yhtä suuriksi. Näin ollen ml-menetelmä on tässä - kuten esim. betajakauman tapauksessa, ks. Vartia (1973, s. 24-26) - tulkittavissa analogiamenetelmäksi:

$$(9) \quad \text{est} f = G \triangleq (x_1 \dots x_n)^{1/n}$$

$$(10) \quad \text{est} E\bar{x} = \bar{x} \triangleq (x_1 + \dots + x_n)/n.$$

Parametrit  $\lambda$  ja  $\mu$  ratkaistaan tämän jälkeen  $f:n$  ja  $E\bar{x}:n$  määritelmäyhtälöistä:

$$(11) \quad \log f = \psi(\mu) - \log \lambda$$

$$(12) \quad E\bar{x} = \mu/\lambda.$$

Nämä yhtälöt voidaan kirjoittaa myös seuraavaan separoitavaan muotoon

$$(13) \quad M = \log \mu - \psi(\mu)$$

$$(14) \quad \underline{Ex} = \mu/\lambda ,$$

jossa  $M$  on suhteellinen (=logaritminen) ero odotusarvon ja geometrisen keskiarvon välillä:

$$(15) \quad M = \log(\underline{Ex}/f) .$$

Suure  $M$  on ei-negatiivinen, koska  $\underline{Ex} \geq f$  ja  $100M$  ilmaisee montako dyniä suurempi  $\underline{Ex}$  on  $f$ :ää, vrt. Vartia (1976, s. 6).

Seuraavaksi johdatetaan eräitä yksinkertaisia  $M$ :n funktioita  $\mu_i = \mu_i(M)$ , jotka approksimoivat tarkasti (13):n yksikäsitteistä ratkaisua  $\mu = \mu(M)$ , parametria  $\mu$  lausuttuna  $M$ :n funktiona. Parametri  $\lambda$  saadaan rekursiivisesti (14):stä lausuttuna  $M$ :n ja  $\underline{Ex}$ :n funktiona  $\lambda = \lambda(M, \underline{Ex})$ . Ml-estimaattorit saavat vastaavasti esitykset  $\hat{\mu} = \mu(\hat{M})$  ja  $\hat{\lambda} = \lambda(\hat{M}, \underline{x})$ , joissa  $\hat{M} = \log(\bar{x}/G)$ .

### 3. Yhtälön (13) likimääräinen ratkaiseminen

Psii-funktiota olen tarkastellut aiemmin, Vartia (1973, s. 88-94), jossa olen mm. johtanut ja taulukoinut seuraavan arvion

$$(16) \quad \psi_3(x) = \log x - \frac{1}{x} \left( \frac{1+3x}{1+6x} \right) .$$

Tämä on kahden merkitsevän numeron tarkkuudella oikea, kun  $0 < x \leq 2$  ja tarkentuvasti oikea, kun  $x \rightarrow \infty$ . Sijoittamalla (16) yhtälöön (13) saadaan

$$(17) \quad M \approx \frac{1}{\mu} \left( \frac{1+3\mu}{1+6\mu} \right) .$$

Peräkkäisten sijoitusten menetelmällä saadaan

$$\begin{aligned} (18) \quad \mu &\approx \frac{1}{M} \left( \frac{1+3\mu}{1+6\mu} \right) \\ &\approx \frac{1}{M} \left( \frac{1+3/2M}{1+6/2M} \right) , \text{ koska } \mu \approx 1/2M \\ &= \frac{1}{M} \left( \frac{2M+3}{2M+6} \right) \\ &= \frac{1}{2M} + \frac{1}{6+2M} . \end{aligned}$$

Tätä  $\mu(M)$ :n arviota merkitsemme  $\mu_1(M)$ :llä

$$(19) \quad \mu_1(M) = \frac{1}{2M} + \frac{1}{6+2M} .$$

Taulukosta 1 havaitsemme, että  $\mu_1(M)$  on kaikkialla melkoisen tarkka  $\mu(M)$ :n approksimaatio sen maksimivirheen ollessa pienempi kuin 0.015. Lisäksi  $\mu_1(M)$ :llä on seuraavat ominaisuudet:

$$(20) \quad \mu_1(M) > \mu(M) \text{ kaikilla } M\text{:n arvoilla,}$$

$$(21) \quad \mu_1(M) \rightarrow \mu(M), \text{ kun } M \rightarrow 0, \text{ jolloin } \mu(M) \rightarrow \infty ,$$

$$(22) \quad \mu_1(M)/\mu(M) \rightarrow 1, \text{ kun } M \rightarrow \infty, \text{ jolloin } \mu(M) \rightarrow 0.$$

Käyttämällä psii-funktion klassisen kehitelmän

$$(23) \quad \psi(x) = \log x - \frac{1}{2x} - \frac{1}{12x^2} + \frac{1}{120x^4} + O\left(\frac{1}{x^6}\right)$$

ensimmäisiä termejä, voidaan peräkkäisten sijoitusten menetelmää ja geometrisen sarjan summakaavaa käyttämällä johtaa seuraavat yhtälöt:

$$(24) \quad \mu(M) = \frac{1}{2M} + \frac{1}{6} - \frac{M}{18} - \frac{M^2}{135} + O(M^3),$$

$$(25) \quad \mu(M) = \frac{1}{2M} + \frac{1}{6+2M} - \frac{M^2}{90} + O(M^3),$$

$$(26) \quad \mu(M) = \frac{1}{2M} + \frac{1}{6+2M+\frac{26}{15}M^2} + O(M^3).$$

Näissä  $O(y)$  merkitsee funktiota, jolle  $O(y)/y$  pysyy rajoitettuna, kun  $y \rightarrow 0$ . Voimme todeta, että kaavojen (25) ja (26) nojalla  $\mu_1(M)$ :n virhe on  $M$ :ssä toista astetta kun  $M \rightarrow 0$  eli  $\mu(M) \rightarrow \infty$ . Taulukosta 1 havaitaan, että seuraava arvio

$$(27) \quad \mu_2(M) = \frac{1}{2M} + \frac{1}{6+3M}$$

on, päinvastoin kuin  $\mu_1(M)$ , jatkuvasti  $\mu(M)$ :n 'aliampuja' eli

$\mu(M) > \mu_2(M)$ . Vaikka  $\mu_2(M) \rightarrow \mu(M)$ , kun  $M \rightarrow 0$  eli  $\mu(M) \rightarrow \infty$ , on se tällöin selvästi  $\mu_1(M)$ :ää heikompi approksimaatio.

Tosin  $\mu_2(M)$ :n maksimivirhe on pienempi kuin 0.01. Näin on esitetty  $\mu(M)$ :lle melkoisen tarkat ylä- ja alarajat:

$$(28) \quad \frac{1}{2M} + \frac{1}{6+2M} > \mu(M) > \frac{1}{2M} + \frac{1}{6+3M}.$$

Todettakoon, että aikaisemmin näyttää tiedetyn ainoastaan,

että  $\mu(M) \approx 1/2M$ , esim Choi & Wette (1969) eivät muuta totea.

Olen maininnut eo. arviot (28) jo aiemmin, Vartia (1972, s. 510 ja 1973, s. 82). Kuviossa 1 on esitetty funktiot  $\mu_1(M)$  ja  $\mu(M)$  suureen  $1/M$  määräämässä koordinaatistossa eräiden karkeiden arvioiden kanssa. Arvio  $\mu_2(M)$  olisi kuviossa saman verran  $\mu(M)$ :n alapuolella kuin  $\mu_1(M)$  on sen yläpuolella.

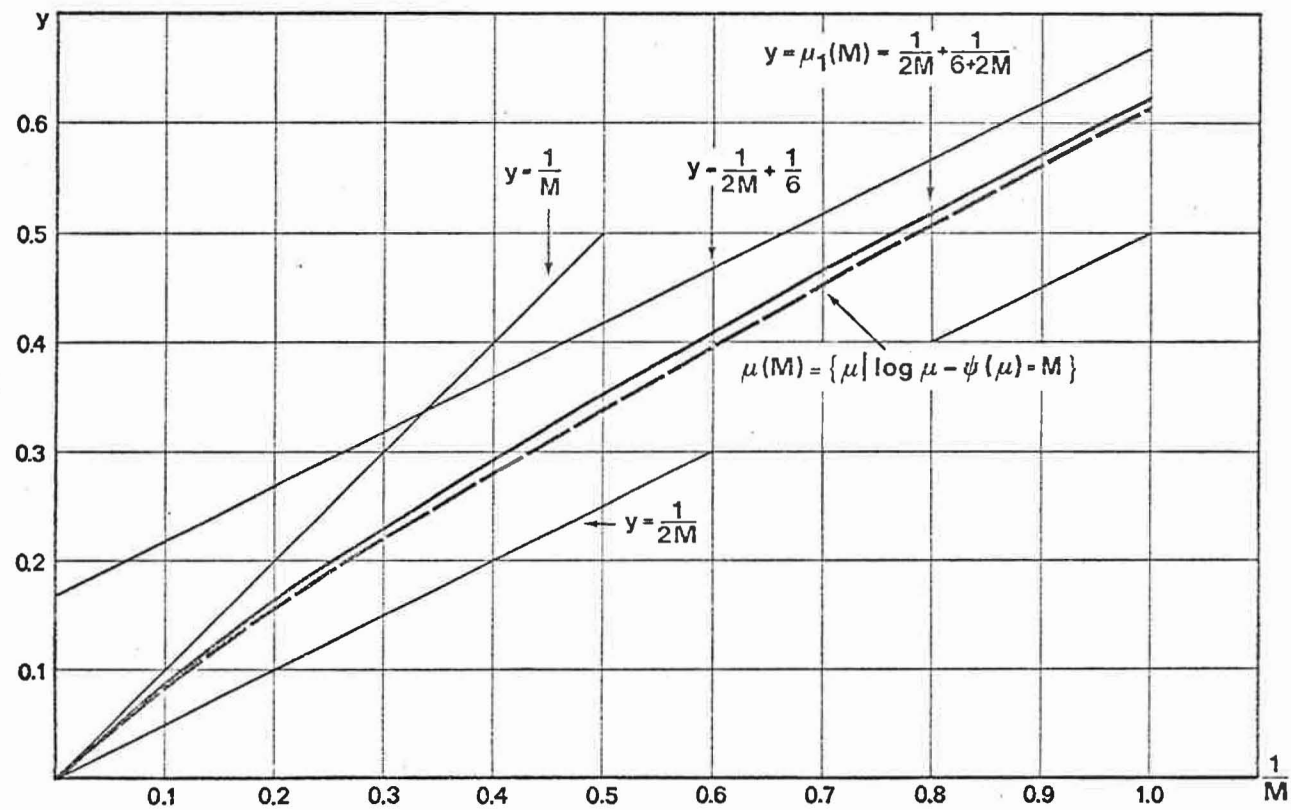
Taulukko 1:  $\mu_1(M)$ ,  $\mu(M)$  ja  $\mu_2(M)$  joillakin  $(1/M)$ :n arvoilla

| $1/M$ | $\mu_1(M)$ | $\mu(M)$  | $\mu_2(M)$ |
|-------|------------|-----------|------------|
| 0.001 | 0.00099985 | 0.0009937 | 0.0008327  |
| 0.005 | 0.004963   | 0.0048839 | 0.004150   |
| 0.02  | 0.0194     | 0.0187155 | 0.0164     |
| 0.10  | 0.0885     | 0.083057  | 0.0778     |
| 0.50  | 0.3500     | 0.338117  | 0.3333     |
| 1.0   | 0.6250     | 0.615557  | 0.6111     |
| 2.0   | 1.1429     | 1.13773   | 1.1333     |
| 3.0   | 1.6500     | 1.64694   | 1.6429     |
| 4.0   | 2.1538     | 2.15185   | 2.1481     |
| 5.0   | 2.6562     | 2.65485   | 2.6515     |
| 6.0   | 3.1579     | 3.15687   | 3.1539     |
| 7.0   | 3.6591     | 3.65830   | 3.6556     |
| 8.0   | 4.1600     | 4.15938   | 4.1569     |
| 9.0   | 4.6607     | 4.66021   | 4.6579     |
| 10.0  | 5.16129    | 5.16088   | 5.15873    |
| 15.0  | 7.663043   | 7.66285   | 7.66129    |
| 20.0  | 10.163934  | 10.16382  | 10.162602  |
| 100.0 | 50.166113  | 50.166108 | 50.165837  |

$\mu_1(M)$ :n ja  $\mu_2(M)$ :n sarakkeisiin on merkitty tarkkuutta osoittavat viivat (oikeat desimaalit).

$\mu(M)$  on laskettu liitteessä esitetyllä iterointimenetelmällä.

Kuvio 1: Funktio  $\mu(M)$  ja eräitä sen approksimaatioita suureen  $(1/M)$  määräämässä koordinaatistossa



4. Tarkkuuskaavat  $\mu(M)$ :lle

Siirrytään konstruoimaan tarkkuuskaavoja  $\mu(M)$ :lle. Kirjoitetaan analogisesti (19):n kanssa

$$(29) \quad \mu(M) = \frac{1}{2M} + \frac{1}{6+2M+\epsilon(M)},$$

jossa funktio  $\epsilon(M)$  määrätään siten, että (29) on identiteetti. Näin muunnetaan  $\mu(M)$ :n approksimointitehtävä sopivasti määritellyn korjaustermiä (tässä  $\epsilon(M)$ :n) arvioimiseksi.

Esimerkiksi korjaustermiä  $\epsilon(M)$  arvioinnissa voidaan sallia melkoisia epätarkkuuksia, koska sen vaikutus yhtälössä (29) on velko vähäinen:  $\mu(M)$ :n pääasiallinen riippuvuus  $M$ :stä sisältyy jo approksimaatioon (19).

Korjaustermiä  $\epsilon(M)$  määrätään siis yhtälöstä

$$(30) \quad \epsilon(M) = 1/[\mu(M) - 1/2M] - (6+2M).$$

Taulukossa 2. on esitetty eräitä  $\epsilon(M)$ :n arvoja. Koska  $\mu_1(M) > \mu(M)$  kaikilla  $M$ :n arvoilla, niin  $\epsilon(M) > 0$  kaikilla  $M$ . Edelleen yhtälön (26) nojalla tiedetään, että

$$(31) \quad \epsilon(M) = \frac{26}{15} M^2 + O(M^3),$$

joten  $\epsilon(M)/M^2 \rightarrow \frac{26}{15}$ , kun  $M \rightarrow 0$ . Tämä kertoo riittävästi  $\epsilon(M)$ :stä pienillä  $M$ :n arvoilla.

Taulukko 2:  $M$ ,  $1/M$ ,  $\mu(M)$  ja  $\epsilon(M)$  joillakin  $M$ :n arvoilla

| $M$  | $1/M$ | $\mu(M)$  | $\epsilon(M)$ |
|------|-------|-----------|---------------|
| 1000 | 0.001 | 0.0009937 | 19.52         |
| 50   | 0.02  | 0.0187155 | 8.738         |
| 10   | 0.10  | 0.083057  | 4.251         |
| 2    | 0.50  | 0.338117  | 1.349         |
| 1.0  | 1.0   | 0.615557  | 0.6537        |
| 0.5  | 2.0   | 1.13773   | 0.2606        |
| 0.2  | 5.0   | 2.65485   | 0.0579        |
| 0.1  | 10.0  | 5.16088   | 0.0158        |

Sensijaan on vaikeampi selvittää kuinka  $\epsilon(M)$  kasvaa  $M$ :n kasvaessa, taulukon 2 mukaan se on joka tapauksessa hyvin hidasta. Muunnetaan yhtälöä (13) käyttämällä psii-funktion siirto-ominaisuutta

$$(32) \quad \psi(x+1) = \psi(x) + \frac{1}{x} ,$$

$$(33) \quad \log \mu - [\psi(\mu+1) - 1/\mu] = M .$$

Edelleen

$$(34) \quad \mu = 1/[M - \log \mu + \psi(\mu+1)] .$$

Koska  $\psi(\mu+1) \rightarrow -\gamma = -0.577216$ , kun  $\mu \rightarrow 0$ , niin voidaan kirjoittaa

$$(35) \quad \psi(\mu+1) - \log \mu = \log \theta M , \text{ jossa}$$

$\theta \rightarrow 1$ , kun  $\mu \rightarrow 0$ . Ja koska  $\mu(M) \cdot M \rightarrow 1$ , kun  $M \rightarrow \infty$  (tai  $\mu \rightarrow 0$ ),

niin suurilla  $M$ :n arvoilla pätee

$$(36) \quad \mu(M) = 1/(M + \log \theta M) ,$$

jossa  $\theta = \theta(M) \rightarrow 1$ , kun  $M \rightarrow \infty$ . Tämän tuloksen avulla voidaan selvittää  $\epsilon(M)$ :n kasvunopeus, kun  $M$  kasvaa. Kirjoitetaan (36) yhtälöä (29) muistuttavaan muotoon lisäämällä ja vähentämällä oikealta puolelta termi  $1/2M$

$$(37) \quad \begin{aligned} \mu(M) &= \frac{1}{2M} + \left( \frac{1}{M + \log \theta M} - \frac{1}{2M} \right) \\ &= \frac{1}{2M} + \frac{1 - (\log \theta M)/M}{2M + 2 \log \theta M} . \end{aligned}$$

Lauseke  $(\log \theta M)/M$  on vielä eliminointava osoittajasta.

Koska  $(\log \theta M)/M \rightarrow 0$ , kun  $M \rightarrow \infty$  voidaan arvioida

$$(38) \quad 1 - (\log \theta M)/M = \frac{1}{1 + (\log \theta' M)/M} ,$$

jossa  $\theta' \rightarrow 1$ , kun  $M \rightarrow \infty$ . Kertomalla nimittäjässä ja yhdistämällä asympotoottisesti samat termit saadaan

$$(39) \quad \mu(M) = \frac{1}{2M} + \frac{1}{2M + 4 \log \bar{\theta} M + 2 (\log \bar{\theta} M)^2 / M} ,$$

jossa  $\bar{\theta} = \bar{\theta}(M) \rightarrow 1$ , kun  $M \rightarrow \infty$ . Viimeinen osoittajan termi häviää, kun  $M \rightarrow \infty$ , mutta muut termit kasvavat rajatta.

Voidaan siis todeta, että  $\mu(M)$  on suurilla  $M$ :n arvoilla seuraavaa muotoa

$$(40) \quad \mu(M) = \frac{1}{2M} + \frac{1}{6 + 2M + 4 \log \tilde{\theta} M} ,$$

jossa  $\tilde{\theta} = \tilde{\theta}(M) \rightarrow 1$ , kun  $M \rightarrow \infty$ . Yhtälöllä (29) määritelty  $\epsilon(M)$  siis kasvaa kuten  $4 \log M$ , kun  $M$  kasvaa rajatta.

Kuitenkin muotoa  $4 \log \tilde{\theta} M$  oleva termi (jos  $\tilde{\theta} \approx 1$ ) käyttäytyy

aivan virheellisesti kun  $M \rightarrow 0$ . Se muuttuu mm. negatiiviseksi, vaikka  $\epsilon(M)$  on aina positiivinen. Sensijaan kaikkialla oikean suuruusluokan antava  $\epsilon(M)$ :n arvio on

$$(41) \quad \epsilon_1(M) = 2\log(1+M^2),$$

sillä  $\epsilon_1(M) \rightarrow 4\log M$ , kun  $M \rightarrow \infty$  ja  $\epsilon_1(M)$  häviää samalla vauhdilla kuin  $M^2$ , kun  $M \rightarrow 0$ .

Vaikka (41) käyttäytyy oikein suurilla ja pienillä  $M$ :n arvoilla antaa se liian suuria arvoja, kun  $M$  on suuruusluokkaa yksi. Eräs hyvin yksinkertainen (vaikkakin suurilla  $M$ :n arvoilla liian hitaasti kasvava) arvio on

$$(42) \quad \epsilon_3(M) = \log(1+M^2).$$

Vastaavaa  $\mu(M)$ :n approksimaatiota merkitään  $\mu_3(M)$ :llä,

$$(43) \quad \mu_3(M) = \frac{1}{2M} + \frac{1}{6 + 2M + \log(1+M^2)}.$$

Arvion  $\mu_3(M)$  maksimivirhe on pienempi kuin  $3 \cdot 10^{-3}$ . Tämän se saa  $M$ :n arvolla 2, jolloin  $\mu(M) \approx 1/3$ . Suhteellinen virhe saavuttaa maksiminsa n. 0.6 % samoilla alueilla. Kun  $M \rightarrow 0$  häviää sekä absoluuttinen että (tietysti) suhteellinen virhe nopeasti. Absoluuttinen virhe häviää samoin kun  $M \rightarrow \infty$ , jolloin suhteellinen virhe pysyy n. 0.5 %:ssa. Arvion  $\mu_3(M)$  tarkkuus ilmenee taulukosta 3. Olen käyttänyt esitettyä  $\mu(M)$ :n approksimaatiota gammajakauman muotoisten betajakaumien maximum-likelihoodyhtälöiden ratkaisemisessa ja saanut sen avulla johdettua yksinkertaiset likimääräiset esitykset betajakauman parametrien  $a$  ja  $b$  ml-estimaattoreille, ks. Vartia (1973, s.82). Muut yritykset johtivat joko huomattavan monimutkaisiin lausekkeisiin tai olivat epätarkkoja.

Taulukko 3:  $1/M$ ,  $\epsilon(M)$ ,  $\epsilon_3(M) = \log(1+M^2)$ ,  $\mu_3(M)$  ja  $\mu_3(M)$ :n suhteellinen virhe eräillä  $M$ :n arvoilla

| $M$  | $1/M$ | $\epsilon(M)$ | $\log(1+M^2)$ | $\mu_3(M)$ | $\frac{\mu_3 - \mu}{\mu}$ , % |
|------|-------|---------------|---------------|------------|-------------------------------|
| 1000 | 0.001 | 19.52         | 13.82         | 0.0009951  | +0.2                          |
| 50   | 0.02  | 8.738         | 7.824         | 0.01879    | +0.4                          |
| 10   | 0.10  | 4.251         | 4.615         | 0.0827     | -0.5                          |
| 2    | 0.50  | 1.349         | 1.609         | 0.3361     | -0.6                          |
| 1.0  | 1.0   | 0.654         | 0.693         | 0.6150     | -0.1                          |
| 0.5  | 2.0   | 0.2606        | 0.223         | 1.1384     | +0.1                          |
| 0.2  | 5.0   | 0.0579        | 0.039         | 2.6553     | +0.02                         |
| 0.1  | 10.0  | 0.0158        | 0.010         | 5.1610     | +0.002                        |

Tutkimalla tarkemmin korjaustermiä  $\epsilon(M)$  käyttäytymistä erityisesti pisteen  $M = 1$  ympäristössä johdin sille seuraavan vielä melko yksinkertaisen approksimaation

$$(44) \quad \epsilon_4(M) = 2 \log(1 + 0.387M^{1.3}).$$

Vastaavan  $\mu(M)$ :n approksimaation

$$(45) \quad \mu_4(M) = \frac{1}{2M} + \frac{1}{6 + 2M + 2 \log(1 + 0.387M^{1.3})}$$

absoluuttinen virhe on alle yhden tuhannesosan ja suhteellinen virhe on pienempi kuin 0.3 %. Tämä arvio on siis noin 2 kertaa tarkempi kuin  $\mu_3(M)$  ja noin 20 kertaa tarkempi kuin  $\mu_1(M)$ . Kun  $M \leq 5$  on  $\mu_4(M)$ :n suhteellinen virhe pienempi kuin 0.1 %. Todettakoon, että  $\mu_4(M)$ :n ja  $\mu(M)$ :n kuvaajien erottamiseksi toisistaan täytyisi kuvio 1 piirtää 10 kertaa suurempaan kokoon.

Liite: Algoritmi yhden muuttujan yhtälön ratkaisemiseksi

Tarkastellaan jatkuvan bijektion  $f$  määrittelemän yhtälön

$$(1) \quad x = f(y)$$

ratkaisemista  $y$ :n suhteen, kun  $x$  on annettu luku  $\hat{x}$ . Tehtävänä on siis laskea  $f(y)$ :n käänteisfunktion  $f^{-1}(x)$  arvo  $\hat{y}$  pisteessä  $\hat{x}$ :

$$(2) \quad \hat{y} = f^{-1}(\hat{x}) .$$

Tämä on usein vaikea tehtävä. Vaikka  $f(y)$  olisi (tietyllä tarkkuudella) helposti laskettava  $y$ :n funktio, kuten esim.  $f(y) = \log y - \psi(y)$ , niin  $f^{-1}(x)$  voi olla hankala  $x$ :n funktio, jota on vaikea riittävällä tarkkuudella lausua tunnettujen funktioiden avulla.

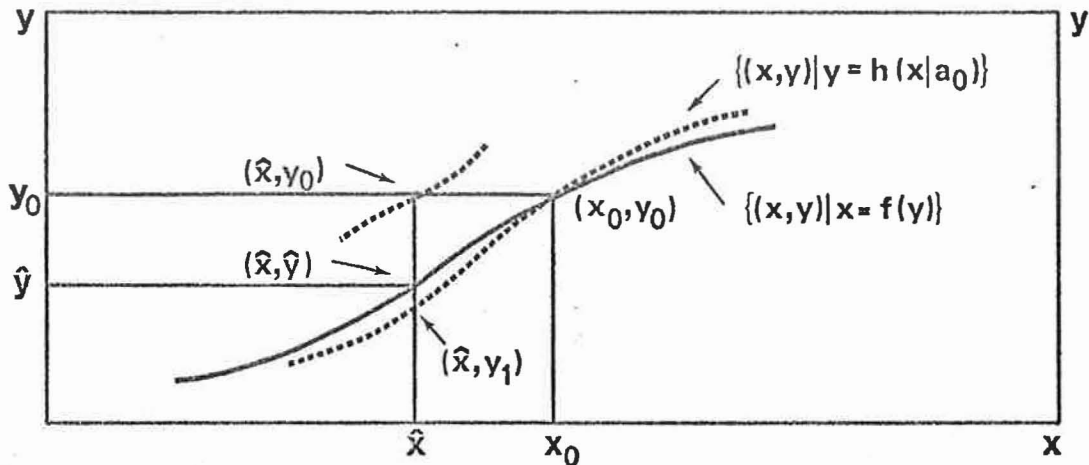
Seuraavassa oletetaan konstruoiduksi parametrissa  $a$  riippuva  $f^{-1}(x)$ :ää approksimoiva funktio

$$(3) \quad h(x|a) \approx f^{-1}(x) ,$$

jolla on seuraavat ominaisuudet.

- A. Kaikilla relevanteilla lukupareilla  $(x_0, y_0)$ , joille  $x_0 = f(y_0)$ , on olemassa parametrin  $a$  arvo  $a_0$  siten, että  $h(x_0|a_0) = y_0$  ja  $h(x|a_0) \approx f^{-1}(x)$ . Tämä merkitsee, että funktiota  $h(x|a_0)$  voidaan käyttää  $f^{-1}(x)$ :n approksimaationa ainakin pisteen  $x_0$  ympäristössä,
- B. Kaikilla relevanteilla lukupareilla  $(x_0, y_0)$ , joille  $x_0 = f(y_0)$  yhtälö  $y_0 = h(x_0|a)$  on helposti ratkaistavissa  $a$ :n suhteen. Tämä merkitsee, että funktio  $h(x|a)$  on helposti sovitettavissa  $f^{-1}(x)$ :n approksimaatioksi mielivaltaisessa pisteessä  $x_0$ .

Liitekuvio: Approksimoitava käänteisfunktio  $f^{-1}(x)$  ja sen approksimaatio  $h(x|a_0)$ .



Olkoon  $y_0 = h(\hat{x})$  yhtälön  $\hat{x} = f(y)$  ratkaisun  $\hat{y}$  ensimmäinen likiarvo, joka on laskettu jonkin  $f^{-1}(x)$ :n approksimaation  $h(x)$  avulla. Tätä alkuarvoa  $y_0$  tarkennetaan seuraavalla algoritmilla.

1. Laske  $x_0 = f(y_0)$  eli  $y_0$ :aa vastaava  $x$ :n arvo.
2. Määrää  $a_0$ 's.e.  $y_0 = h(x_0|a_0)$  eli käyrä  $y = h(x|a_0)$  kulkee pisteen  $(x_0, y_0)$  kautta.
3. Laske  $y_1 = h(\hat{x}|a_0)$ . Tämä on ratkaisun  $\hat{y} = f^{-1}(\hat{x})$  uusi likiarvo.
4. Palaa kohtaan 1. ja käytä lukua  $y_1$  uutena alkuarvona  $y_0$ :n sijalla.

Jos verrataan edellä lyhyesti luonnehdittua menetelmää tavallisiin yhtälön ratkaisemisessa sovellettuihin iterointimenetelmiin kuten Newtonin menetelmään, Mullerin menetelmään (ns. Regula falsi) tai tavalliseen iterointiin (ks. Henrici (1964)) niin on todettava, että  $f^{-1}(x)$ :ää approksimoivat funktiot  $h(x|a)$  voivat edellä olla epälineaarisia ja globaalisia  $f^{-1}(x)$ :n approksimaatioita. Mainituissa tavanomaisissa menetelmissä  $h(x|a)$ :t ovat lineaarisia  $x$ :n funktioita ja toimivat vain lokaalisina  $f^{-1}(x)$ :n approksimaatioina. Pyrkimyksenä on tässä ollut globaalin approksimoinnin ideoiden hyväksikäyttö yhtälön  $x = f(y)$  ratkaisemisessa: funktiot  $h(x|a_n)$  voivat olla yleisemminkin mielenkiintoisia.

Sovelletaan menettelyä yhtälön  $x = \log y - \psi(y)$  ratkaisemiseen  $y$ :n suhteen kun  $x = \hat{x}$ . Tässä siis  $f(y) = \log y - \psi(y)$ . Ensimmäisenä  $\hat{y}$ :n likiarvona käytetään lukua

$$(4) \quad y_0 = \frac{1}{2\hat{x}} + \frac{1}{6+2\hat{x}} = h(\hat{x}) .$$

Käänteisfunktion  $f^{-1}(x)$  approksimaationa olkoon

$$(5) \quad h(x|a) = \frac{1}{2x} + \frac{1}{6+2x+a\log(1+x^2)} .$$

Lasketaan kohdan 1. mukaisesti  $x_0 = f(y_0) = \log y_0 - \psi(y_0)$ .

Määritetään tämän jälkeen  $a_0$  yhtälöstä

$$(6) \quad y_0 = \frac{1}{2x_0} + \frac{1}{6+2x_0+a_0\log(1+x_0^2)} .$$

$$(7) \quad a_0 = \left[ \frac{1}{(y_0 - \frac{1}{2x_0})} - (6+2x_0) \right] / \log(1+x_0^2) .$$

Saatu funktio  $y = h(x|a_0)$  kulkee pisteen  $(x_0, y_0)$  kautta.

Tarkennettu  $\hat{y}$ :n arvio on

$$(8) \quad y_1 = h(\hat{x}|a_0) = \frac{1}{2\hat{x}} + \frac{1}{6+2\hat{x}+a_0 \log(1+\hat{x}^2)}.$$

Esitetty iterointikaavio on parilla kymmenellä Basic-käskyllä ohjelmoitavissa ja se konvergoi erittäin nopeasti. Kuuden desimaalin tarkkuuteen tarvittiin korkeintaan kolme iterointikierrosta. Taulukon 1 ratkaisupisteet on laskettu GE 600:lla tietokonepäättteen avulla.

## Lähdeviitteet

ABRAMOWITZ, M., & STEGUN, I.A., (1970): "Handbook of mathematical functions with formulas, graphs, and mathematical tables", Government Printing Office, USA.

CHOI, S.C., & WETTE, R., (1969): "Maximum likelihood estimation of the parameters of the gamma distribution and their bias", Technometrics, 11, 683-690.

CRAMÉR, H., (1946): "Mathematical methods of statistics", Princeton University Press, USA.

HENRICI, P., (1964): "Elements of numerical analysis", John Wiley & Sons, Inc., New York.

MC DONALD, J.B., & JENSEN, B., (1975): "An analysis of some properties of alternative measures of income inequality based upon the gamma distribution function", Department of Economics, Brigham Young University, USA (esitetty Econometric Societyn Helsingin kokouksessa v. 1976).

RAO, C.R., (1968): "Linear statistical inference", John Wiley & Sons, Inc., New York.

SALEM, A.B., & MOUNT, T.D., (1974): "A convenient descriptive model of income distribution: the gamma density", Econometrica, 42, 1115-1127.

VARTIA, Y.O., (1972): "Johdatus tilastotieteeseen, osa II", Gaudeamus, Helsinki.

VARTIA, Y.O., (1973): "Betajakauman parametrien estimointi maximum likelihood-menetelmällä", Elinkeinoelämän Tutkimuslaitoksen julkaisuja A1, Helsinki.

VARTIA, Y.O., (1976): "Relative changes and index numbers", The Research Institute of the Finnish Economy, Serie A4, Helsinki.