

# GALLUPEISTA, VIRHEMARGINAALISTA JA VAALITAVASTA

JUHA ALHO – ERKKI PAHKINEN – MIKKO A. SALO

”Aho ja Halonen kamppailevat tasoissa” otsikoi Helsingin Sanomat (3.2.2000) juuri ennen presidentinvaalipäivää. Tieto perustui 1.2. tehtyyn puhelinhaastatteluun, johon vastasi 1 435 henkilöä. Halosen kannatuksiksi saatiin 51 prosenttia ja Ahon 49 prosenttia. Itse tekstissä todettiin, että ero mahtui selvästi tutkimuksen virhemarginaaliin  $\pm 3$  prosenttia. Vaalien tulos oli 51,6 prosenttia Haloselle ja 48,4 prosenttia Aholle.

Ero otsikoinnissa, tutkimuksen teknisessä kuvauksessa sekä tulosten analyysissä verrattuna edellisiin presidentinvaaleihin on merkittävä. Tuolloin haastattelututkimusten tuloksia raportoitiin jopa desimaalien eroja painottaen. Ilahtuneena olemme todenneet, että myös radiossa ja TV:ssä virhemarginaalit otetaan nykyään säännöllisesti huomioon gallupien tuloksia arvioitaessa.

Pertti Suhonen (1999) tarkasteli runsas vuosi sitten puolueiden kannatusmittausten tarkkuutta ansiokkaassa artikkelissa, jossa tuotiin esiin – tietääksemme ensimmäistä kertaa – teknisiä yksityiskohtia kaupallisten tutkimuslaitosten käyttämistä estimointimenetelmistä. Hyödyllisestä sisällöstään huolimatta Suhosen artikkeli päättyi liian jyrkkiin johtopäätöksiin. Sen mukaan virhemarginaalit pikemminkin ”johtavat harhaan ja vaikeuttavat luotettavuuden arvioimista”. Edelleen Suhonen toistaa usein esitetyn argumentin, että koska virhemarginaali kuvastaa vain otantasattumaa, aineistonkeruun menetelmistä ja kannatusarvioiden laskutavasta aiheutuvat tärkeämmät virheet jäävät kuvaamatta. Tässäkään ei mielestämme ole syytä näin suoraviivaiseen tulkintaan. Tilastotieteen perusoppikirjoja (esim. Freedman & al. 1978, Chapter 19) myöten on pitkään tiedetty, että monet seikat vaikuttavat tulosten tarkkuuteen. Eri virhelähteiden suuruusluokat vaihtelevat kuitenkin voimakkaasti tutkimukses-

ta ja maasta toiseen, eikä otantasattumaa voi jättää huomioon ottamatta (Groves 1989).

Vaalien yhteydessä on käyty myös vilkasta keskustelua sekä gallupien vaikutuksesta äänestystulokseen että kaksivaiheisen valintatavan rakaudesta.

Koska kaikki edellä mainitut teemat vaikuttavat siihen, miten demokratia toteutuu maassamme, pyrimme seuraavassa kuvaamaan gallupien ja kaksivaiheisen vaalin yhdessä muodostamaa valintamenetelmää. Käsittääksemme kahdet viimeksi käydyt presidentinvaalit osoittavat, että menetelmä toimii kohtuullisesti. Kummallakin kerralla loppusuoralle selvisi kaksi pätevää ehdokasta, toisin kuin monet politiikan tutkijat olivat pelänneet. Silti on hyvä huomata, että järjestelmää voisi yksinkertaistaa.

## TODENNÄKÖISYYSSOTANTA JA VALIKOIDUT OTOKSET

Katsaus taaksepäin osoittaa, että kyselytutkimusten pääasialliset virhelähteet ovat jo kauan olleet tunnettuja. Vuoden 1936 presidentinvaaleissa Yhdysvalloissa olivat vastakkain silloinen presidentti Roosevelt sekä republikaanien puolelta Kansasin kuvernööri Landon. Rooseveltiä pidettiin suosikkina, mutta arvostettu Literary Digest -lehti ennusti, että Landon voittaisi 57 prosenttia äänistä Rooseveltin 43:a prosenttia vastaan. Ennustus perustui tietävästi suurimpaan koskaan järjestettyyn kyselytutkimukseen (Freedman & al. 1978). Vastaajia oli 2,4 miljoonaa. Vaaleissa Roosevelt sai 62 prosenttia äänistä. George Gallupille, joka oli tuolloin aloittamassa survey-tutkimusten tekoa, sen enempää vaalin tulos kuin Literary Digestin ennustukseen eivät tulleet yllätyksenä. Ennen kummankaan tuloksen selviä-

mistä hän ennusti 3 000 henkilön otoksen perusteella, että a) Literary Digest tulisi ennustamaan Rooseveltille 44:ää prosenttia äänistä ja b) että Roosevelt tulisi saamaan 56 prosenttia äänistä.

Literary Digestin haastattelututkimuksen epäonnistumiselle oli kaksi pääsyötä (Freedman & al. 1978). Ensinnäkin, *otannan kehikko* ei vastannut koko äänestäjäkuntaa, sillä kyselytutkimuksen osoitetiedot perustuivat puhelinluetteloihin ja erilaisten järjestöjen jäsenluetteloihin. Hyvätulaiset olivat kehikossa ylliedustettuina. Toiseksi, kysely oli lähetetty 10 miljoonalle henkilölle ja vain 24 prosenttia vastasi. *Kato* oli siis 76 prosenttia. Yleisesti on tunnettua, että köyhät ja kaikkein rikkaimmat ovat haluttomimpia vastaamaan kyselyihin. Jälkimmäisiä on vähän, joten kato kasvatti keinotekoisesti Landonin kannattajien osuutta otoksessa.

George Gallupin menetelmän vahvuutena oli todennäköisyysteoriaan perustuva otantamenetelmä. Tällöin kannatusarvion tarkkuuskin voidaan estimoida käytetyn valintamenetelmän perusteella. Yksinkertaisen satunnaisotannan tapauksessa laskenta perustuu ns. hypergeometrisen jakauman teoriaan. Perustulokset ovat vanhoja ja siinä määrin yksinkertaisia, että ne nykyisin esitetään tilastotieteen peruskursseilla yliopistoissa. George Gallupin merkittävin ansio olikin käytännöllinen. Hän osoitti, että kyselytutkimukset voidaan toteuttaa kohtuullisin kustannuksin ja niillä saavutetaan tarkempia tuloksia kuin sellaisilla harhaisilla menetelmillä, joita myös Literary Digest käytti.

Se, että tieteellistä lähestymistapaa mielipiteiden kartoittamiseen ei aina ole ollut, unohtuu monilta gallupitutkimusten kriitikoilta. Asia pyrkii hämärtyämään siitäkin syystä, että viime aikoina radioon ja TV:hen on ilmaantunut erilaisia kuuntelija- tai katsojajäännöstyksiä, joissa puhelimitse tai Internetin kautta voi ilmoittaa mielipiteensä eri asioista. Näiltä kyselyiltä puuttuu kokonaan gallupitutkimusten tieteellinen pohja, ja niiden osuvuus tuskin on parempi kuin Literary Digestin kyselyn.

George Gallupin otantamenetelmä ei sekään ollut vailla ongelmia. Tämä tuli ilmi Yhdysvalloissa vuoden 1948 presidentinvaalissa, jossa olivat vastakkain Truman ja Dewey. Gallup arvioi Deweyn kannatukseksi 50 prosenttia ja Trumanin 44 prosenttia. Itse vaalissa prosentit olivat päinvastoin: 50 % Trumanille ja 45 % Deweylle. Syy-

nä oli Gallupin käyttämä *kiintiöpoiminta* (Freedman & al. 1978). Tässä menetelmässä otantaa jatketaan niin kauan, että ennalta valittujen taustamuuttujien (esim. ikä, sukupuoli ja asuinpaikka) mukaisiin luokkiin tulee otoksessa sama osuus kuin perusjoukossakin on. Riippuen haastattelijalle annetuista ohjeista tämä menettely voi olla lähellä todennäköisyyslaskentaan perustuvaa ns. ositettua otantaa tai sitten se voi olla lähellä harkintaotantaa (Cochran 1977, 135–136). Erityisesti haastattelusta kieltäytyminen voi aiheuttaa harhaa, jos kieltäytymisen todennäköisyys riippuu vastaajan poliittisesta kannasta. Vuoden 1948 vaalien aikaan haastattelijoilla oli melko suuri valta valita haastateltavansa, ja tämä ilmeisesti aiheutti virheellisiä tuloksia (Freedman & al. 1978).

Suomessakin käytetään kiintiöpoimintaa kannatusosuuksien arvioinnissa. Otos poimitaan niiden väestörekisterissä olevien joukosta, joilla on puhelin. Ne, joita ei tavoiteta tai jotka kieltäytyvät vastaamasta, korvataan kiintiöntikriteerien mukaisilla henkilöillä. Tämä muistuttaa tekniikkaa, jota käytetään puuttuvien tietojen paikkaukseen puhtaassa todennäköisyysotannassa. Menetelmän harhattomuus riippuu kuitenkin kadon syistä ja toteutuksen yksityiskohdista (Rubin 1987).

#### OTANTAMENETELMÄ JA VIRHEMARGINAALI

Suhosen (1999) artikkelin yksi ansio on, että siinä kuvataan suomalaisten survey-organisaatioiden otantamenetelmiä. Kyseisissä menetelmissä käytetään tietoa aiemmasta äänestämisestä aputietona, jonka avulla voidaan kuvata muutoksia puolueiden kannatuksissa. Näillä tiedoilla on merkitystä myös presidentinvaaleissa ehdokkaiden puoluetustaan takia. Tästä huolimatta organisaatiot käyttävät yksinkertaisen satunnaisotannan kaavoja keskivirheiden laskennassa (vrt. Suhonen 1999, 67). Sikäli kuin survey-organisaatioiden käyttämät otanta- ja estimointimenetelmät muistuttavat ositettua otantaa, niiden voidaan olettaa olevan tarkempia kuin yksinkertainen satunnaisotanta. Kiintiöpoiminta aiheuttaa kuitenkin uuden virhelähteen, vaikka sen vaikutuksia yritetäänkin korjata painottamalla. Käytetyt kaavat lienevät siten oikean suuntaisia approksimaatioita. Konservatiivisen arvion (ts. liian suuren) käyttö olisi sikäli perusteltua, että mah-

dollisista kehikko- tai katovirheistä aiheutuvaa epävarmuutta voidaan sen avulla osin ottaa huomioon.

Suomessa väestökisteri tarjoaa varsin luotettavan kehikon otannalle, ja vaikka kato on kaikissa haastattelututkimuksissa potentiaalinen ongelma, esimerkiksi Suhosen (1999, 67) mainitsemassa eduskuntavaaligallupissa se näyttäisi olleen 18 prosenttia. Edellytykset melko luotettavaan estimointiin ovat Suomessa näiltä osin siis olemassa. Kuluttajansuojan takia olisi kuitenkin tärkeää, että survey-organisaatiot esittelisivät käyttämäänsä menetelmiä jollain julkisella foorumilla. Olisi erityisen kiinnostavaa tietää, miten kehikkotai katovirheet ja otannan kiintiöiminen vaikuttavat tarkkuuteen Suomessa.

#### GALLUPIEN VAIKUTUS ÄÄNESTYSKÄYTTÄYTYMISEEN

Viime presidentinvaaleille oli tunnusomaista usean ehdokkaan gallup-suosiossa tapahtuneet jyrkät muutokset vaalien alla. Eräät tutkijat ovat esittäneet, että gallup tulosten esittämistä vaalien alla pitäisi rajoittaa. Käsittääksemme gallupien luonne on ymmärretty tällöin väärin. Onhan jo todettu, että kyselyyn vastaaminen ja äänestäminen ovat eri asioita (Freedman & al. 1978). Kyselyssä pyydetään vastausta epätodeksi tiedetyn hypoteesin pohjalta ("Jos vaalit olisivat tänään, ketä äänestäisitte?"). Esimerkiksi se, keitä vastaushetkellä pidetään "varteenotettavina ehdokkaina" voi merkittävästi vaikuttaa vastaajan kantaa. Näin ollen gallupeihin liittyy volatilitteettia, jota itse vaalitulanteessa ei samalla tavalla ole.

Tällainen mielipiteiden vaihtelu vaikuttaa gallupin tulokseen kuitenkin vain, jos se tapahtuu samanlaisena useiden äänestäjien kohdalla. Jos mielipiteen ailahtukset tapahtuisivat toisistaan riippumatta eri kannattajien hyväksi, niiden vaikutus olisi (ns. suurten lukujen lakien perusteella, ks. Chung 1974, Chapter 5) vaatimaton. Gallupit mittaavat siis suurissa ihmisjoukoissa tapahtuvaa samanaikaista mielipiteiden ailahtelua. Näiden "sosiaalisten faktojen" tekeminen julkiseksi on merkittävä demokraattisen päätöksenteon piirre. Julkisten gallup tulosten pohjalta äänestäjillä on mahdollisuus itse muodostaa kantansa ja äänestää joko virran mukaan tai sitä vastaan.

Onnetonta olisi, jos gallup tulosten julkaisua rajoitettaisiin. Koska ihmiset tietävät mielipitei-

den ailahtelevan, tietoa näistä muutoksista tul-taisiin hakemaan muilla keinoin. Toimittajista, puoluevirkaileijoista, tutkijoista ja muista politiikan kommentaattoreista tulisi oraakkeleita, joita pyydetäisiin arvailemaan muutosten suuntaa. Tämä kaventaisi demokratiaa ja kasvattaisi poliittisen eliitin valtaa.

Kysymys siitä, vaikuttaako itse gallup tuloksen julkaiseminen äänestäjien valintoihin, on avoin. Vaikka ihmiset ovatkin kiinnostuneita muiden mielipiteistä, ei ole ollenkaan selvää, mihin suuntaan esimerkiksi aiempaa paremman kannatusluvun julkaiseminen vaikuttaa ehdokkaan kannatukseen. Manipulaation mahdollisuuksia ei juuri näyttäisi olevan. Silti asiaa sietäisi tutkia empiirisesti.

Koska mielipidetiedustelu koskee eri ajankoh-taa kuin itse vaali ja on edellä kuvatuista syistä muutenkin enemmän altis vaihtelulle kuin vaalikannatus, on keskeistä, että tämän vaihtelun pääl-le tuleva, otantasattumasta aiheutuva vaihtelu tu-lee oikein kuvatuksi. Jos gallup tulosten pelätään vaikuttavan vaalituloksiin epätoivottavalla tavalla, niin yksinkertainen tapa vähentää tätä vaikutus-ta on vaatia, että kyselyn virhemarginaali rapo-roidaan tarkasti. Tällöin muutosten analyysiin tu-lee automaattisesti mukaan juuri sellaista suur-piirteisyyttä, jota alussa siteeratussa lehtiotsikos-sakin oli.

#### BORDAN ÄÄNESTYSMENETELMÄ JA PRESIDENTINVAALIT

Kaksivaiheisuus pidentää presidentinvaalia ja voi kysyä, olisiko jotain tapaa valita presidentti vai-vaamalla äänestäjiä vähemmän.

Ranskalainen matemaatikko Jean-Charles Bor-da esitti vuonna 1770 vaalijärjestelmän, jossa ää- nestäjä panee ehdokkaat paremmuusjärjestykseen (McLean & Urken 1995; Nurmi 1999, 11–15). Kun ehdokkaita on k kappaletta, paras saa äänes-täjältä k ääntä, toiseksi paras k – 1 ääntä jne. Kaikkiaan eniten ääniä saanut voittaa. Viime vuo-sikymmenellä on pystytty osoittamaan ns. kaaos-teoriaa soveltaen (Saari 1995), että vaikka Bordan järjestelmäkin voi tuottaa paradoksaalisia loppu-tuloksia, niiden mahdollisuus on pienempi kuin muissa ehdotetuissa äänestysmenetelmissä. Tämä on merkittävä matemaattinen tulos, sillä aiem-min äänestysmenetelmiä on voitu perustella vain yksittäisillä paradokseilla, joiden perusteella on

katsottu milloin minkin äänestysmenetelmän olevan käyttökelpoinen.<sup>1</sup>

Tulos perustuu siihen, että Bordan menetelmässä mitataan tavallaan ehdokkaiden absoluuttista hyvyttä järjestyslukujen kautta. Kuten viime presidentinvaaleissa todettiin, uusien ehdokkaiden tulo tai vanhojen poisjäänti voi muuttaa muiden suhteellista kannatusta. Näin käy erityisen helposti enemmistövaalissa, sillä esimerkiksi silloin, kun uusi henkilö tunnustetaan varteenotettavaksi ehdokkaaksi, hänen kannattajiensa osalta tieto muiden ehdokkaiden keskinäisistä suhteista katoaa kokonaan. Muiden ehdokkaiden suhteellisessa kannatuksessa tapahtuvat muutokset eivät siis välttämättä johdu siitä, että äänestäjät olisivat epäjohdonmukaisia preferensseissään, tai siitä, että muiden ehdokkaiden

---

<sup>1</sup> Esimerkki paradoksista. Ehdokkaita ovat A, B ja C ja "A > B" tarkoittaa, että A on parempi kuin B. Oletetaan, että ensimmäisellä kierroksella 40 % äänestäjistä on sitä mieltä, että "A > B > C", 35 % on sitä mieltä, että "C > B > A" ja 25 % on sitä mieltä, että "B > C > A". Toiselle kierrokselle pääsevät enemmistövaalissa A ja C, joista C voittaa saamalla 60 % äänistä A:n 40:tä % vastaan. Huomattakoon, että tässä tapauksessa ensimmäisellä kierroksen enemmistövaalissa huonoimmin menestyvän ehdokkaan B kannatus olisi parempi kuin A:n, 60 % vastaan 40 %. Kenties vielä yllyttävämpää on se, että B:n kannatus on myös suurempi kuin vaalit voittaneen C:n, 65 % vastaan 35 %. Bordan menetelmän mukaan jokaisella äänestäjällä olisi tässä tapauksessa 1 + 2 + 3 = 6 ääntä ja vaaleissa B saisi keskimäärin 2,25 ääntä/äänestäjä, C saisi 1,95 ja A saisi 1,80. Vaalit voitettaisiin siis B.

#### KIRJALLISUUS

- Chung, K. L.: A course in probability theory, 2nd ed. New York: Academic Press, 1974  
Cochran, W. G.: Sampling theory, 3rd ed. New York: Wiley, 1977  
Freedman, D. & Pisani, R. & Purves, R.: Statistics. New York: Norton, 1978  
Groves, R. M.: Survey errors and survey costs. New York: Wiley, 1989  
McLean, I. & Urken, A. (eds.): Classics of social

kannatus suhteessa toisiinsa olisi todellisuudessa muuttunut.

Juuri tästä oli kysymys presidentivaalien toisella kierroksella, jolloin arvailun kohteena oli Rehiniä ja Uosukaista ensimmäisellä kierroksella äänestäneiden käyttäytyminen. Tämä olisi Bordan menetelmää käyttäen tullut selväksi yhdelläkin vaalikierroksella. Erityisesti sellaisissa äänestystilanteissa, joissa useat ehdokkaat vetoavat moniin äänestäjiin, Bordan menetelmä ottaisi enemmistövaalia paremmin huomioon äänestäjien mieltymykset. Samalla vaaleja edeltävien gallupmittausten tulokset stabiloituisivat, sillä uusien ehdokkaiden tulo, vanhojen luopuminen tai yksittäisen ehdokkaan suosion muutokset vaikuttaisivat suhteellisesti vähemmän muiden keskimääräisiin järjestyslukuihin.

Bordan menetelmää on pidetty teknisesti vaikeana toteuttaa. Lienee kuitenkin vain ajan kysymys, milloin siirrytään tietokoneiden käyttöön äänestyksessä. Ongelmallisempaa olisikin kenties se, ovatko äänestäjät halukkaita järjestämään koko ehdokasjoukkoa, joka voi olla suuri. Tämä lienee kuitenkin tottumuskysymys.

Bordan menetelmän käyttöönotto muuttaisi gallupien käytännön toteutusta. Virhemarginaalit eivät enää koskisi kannatusosuutta vaan ehdokkaan saamaa keskimääräistä äänimäärää per äänestäjä. Muutokset olisivat kuitenkin lähinnä teknisiä.

Bordan äänestysjärjestelmän käyttöönotto kannattaisi harkita jo senkin takia, että vaaleja edeltävä kuva kansalaisten preferensseistä tulisi gallupeissa nykyistä luotettavammin esiin.

- choice. Ann Arbor: University of Michigan Press, 1995  
Nurmi, H.: Voting paradoxes and how to deal with them. Berlin: Springer, 1999  
Rubin, D.: Multiple imputation for nonresponse in surveys. New York: Wiley, 1987  
Saari, D. G.: A chaotic exploration of aggregation paradoxes. SIAM Review 37 (1995), 37–52  
Suhonen, P.: Onko puolueiden kannatusmittauksiin uskomista? Yhteiskuntapolitiikka 64 (1999): 1, 62–68.