

**This is an electronic reprint of the original article.  
This reprint *may differ* from the original in pagination and typographic detail.**

**Author(s):** Högmander, Harri

**Title:** Mikä virhemarginaali?

**Year:** 2015

**Version:**

**Please cite the original version:**

Högmander, H. (2015, 26.05.2015). Mikä virhemarginaali?. Tiedeblogi. Retrieved from <https://www.jyu.fi/blogit/tiedeblogi/hogmander>

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

# Mikä virhemarginaali?



Harri Högmander, kuvaaja Enja Heikkilä

Kevään eduskuntavaalien voittajista riittää mielipiteitä. Minä pidän kollegani lausumasta ”tilastotiede voitti”. Tällä hän tarkoitti sitä, että kaikkien puolueiden kannatuslukemat osuivat nätisti tutkimuslaitosten ennusteiden virhemarginaaleihin.

Mikä on gallup-uutisissa esiintyvä virhemarginaali, *margin of error*? Selvästi tämä tunnusluku varsinaisen kannatusosuusarvion lisänä kertoo estimoinnin tarkkuudesta. Siis siitä epävarmuudesta, joka syntyy, kun parin tuhannen vastauksen pohjalta tehdyt laskelmat yleistetään tilastollisen mallin perusteella koskemaan koko äänestäjäkuntaa. (Tuloksen tarkkuus ei muuten riipu siitä, onko äänestäjiä satatuhatta tai miljardeja, mikä on kiehtova matemaattinen fakta sinänsä.)

Tällainen numeeriseen otokseen perustuva laskenta ja yleistämisestä koituvan epätarkkuuden arviointi on tilastotieteen perustehtävä: tunnettua pientä yleistetään tutkittavaan isoon, ja otannasta johtuvaa epävarmuutta hallitaan todennäköisyyslaskennan keinoin. Tilastollisen tutkimuksen tuloksiin jää aina satunnaisuuden seurauksena välys. Tämä hinta maksetaan siitä helpotuksesta, että empiirisen alan tutkijalle riittää hankkia vain pieni otos koko kohteena olevasta todellisuudesta.

Virhemarginaali on tarkkaan ottaen 95 %:n luottamusvälin pituuden puolikas. Tämä tarkoittaa, että satunnainen, samankokoisesta otoksesta laskettu väli sisältää todellisen arvon 95 %:n todennäköisyydellä (tai, 95 % lasketuista väleistä kattaa oikean kannatusosuuden). Todellinen lukema ei siis ole virhemarginaalin rajoissa täysin varmasti – karkeasti keskimäärin joka kahdeskymmenes arvio on huti. Valittu viiden prosentin virhemahdollisuus, joka on vakiintunut empiiriseen tutkimukseen yleisesti, on tilastotieteen suurmiehen R. A. Fisherin idea vajaan sadan vuoden takaa.

Kaikki virhemarginaalin sisältämät lukemat eivät ole teoreettisesti ”yhtä hyviä arvauksia”, vaan luottamusvälin laidoilla olevat luvut ovat vähemmän uskottavia. Esim. jos puolueen kannatuslukemaksi saadaan 20,5 % ja virhemarginaaliksi lasketaan 2,5 prosenttiyksikköä, niin on huomattavasti varmempaa, että todellinen osuus on välillä 20–21 % kuin välillä 22–23 %.

Kannatusosuuden arvioinnissa käytettävä malli nojaa ositettuun otantaan, jossa meidät äänestäjät on jaettu poliittisen mielipiteen vaihtelua pienentäviin osajoukkoihin esim. sukupuolen, iän, asuinpaikan ja ammatin mukaan. Osittaminen tekee otannasta tehokkaampaa kuin täysin satunnainen arvonta. Haastateltavalta kysytään myös, mitä puoluetta hän äänesti edellisissä vaaleissa. Näin voidaan pienentää otannan harhaisuutta, sillä edellisvaalien tulos tiedetään. Jos otokseen sattuu tulemaan suhteessa liian vähän jonkin puolueen äänestäjiä tai nämä eivät jostain syystä kerro kantaansa, voidaan poikkeamaa oikaista laskennallisesti.

Entä miten tilastotieteilijä laskee sen virhemarginaalin? Tekniikka perustuu ns. keskeiseen raja-arvolauseeseen, jonka mukaan otoksista lasketut keskiarvot, mitä gallupien kannatusosuudetkin ovat,

noudattavat normaalijakaumaa. Tämä kaunis, P.-S. Laplacen ja C. F. Gaussin n. v. 1800 todistama tulos on ihmiskunnan tärkeimpiä saavutuksia. Mutta se on jo toinen juttu.

*Harri Högmänder, lehtori, matematiikan ja tilastotieteen laitos.*