

## Tilastollisen päättelyn heikkouden merkit

KETOKIVI, MIKKO (2015). *Tilastollinen päättely ja tieteellinen argumentointi*. 2. uud. p. 336 s. Gaudeamus, Helsinki.

Olen aina halunnut aloittaa tai lopettaa kirja-arvion legendarisen, nyt jo edesmenneen, näyttelijän Simo Salmisen sanoin: ”Dorkaa ja karmeeta, *two points!*” (elokuvasta *Hei kliffaa hei*, 1985). Useaan tilanteeseen sopiva sutkaus ei kuitenkaan kuvaa tuntemuksiani liikeloustieteen professori Mikko Ketokiven teoksen suhteen, ennemminkin päinvastoin. Teosta voidaan pitää ansiokkaana johdantona tilastolliseen päättelyyn sekä hyvänä muistutuksena empiirisen tutkimuksen heikkouksista ja sudenkuopista. Toisin kuin useimmat yksilöteiset suoraviivaisista ohjeista koostuvat menetelmäoppaat ja käsikirjat, Ketokiven teoksen toinen, laajennettu painos kattaa monipuolisesti empiirisen tutkimuksen elinkaaren (huonoine ja hyvine esimerkkeineen) aina tutkimuksen motivoinnista ja kysymysten asetelusta tulosten tulkintaan ja tuloksiin perustuvaan argumentoin-

tiin saakka. Välissä perehdytään laajasti ilmiöiden operationalisointiin, mittaamiseen ja analysointiin.

Ketokivi kehottaa tutkijoita asettamaan rimansa (epämiellyttävän) korkealle luopumalla tilastollista päättelyä leimaavista nyrkkisäännöistä. Kehotus on osuva, sillä liian usein tutkijat (allekirjoittanut mukaan lukien) etsivät tilasto-ohjelmien tuottamista tulosteista aikaisempaan kirjallisuuteen pohjautuvia raja-arvoja – ja ”tuomitsevat” tulokset näiden nyrkkisääntöjen mukaisesti. Jos tulosteista löytyy raja-arvoja ylittäviä tai alittavia numerosarjoja, jotka vahvistavat tutkijan käsitystä ilmiöstä, siirrytään tyytyväisenä tulosten tulkintaan pohtimatta sen tarkemmin oman aineiston tai itse menetelmien sopivuutta, laskuperusteita, heikkouksia tai taustaoletuksia. Onneksi näiden nyrkkisääntöjen välttämiseen löytyy Ketokiven teoksesta nyrkkisääntö: älä koskaan käytä menetelmäkirjallisuudessa tarjottua ratkaisua tai sääntöä ellet ymmärrä kontekstia, jossa sääntö on esitetty, ja oletuksia joihin sääntö perustuu. Tai hieman toisin muotoiltuna: tiettyjä asioita pidetään it-

sestään selvinä, mutta kuten Ketokivi toteaa, kun kyse on tilastotieteestä ja erityisesti empiirisestä tilastollisesta tutkimuksesta, mikään ei ole itsestään selvää.

Teos siis alleviivaa ansiokkaasti suuren määrän tilastolliseen empiiriseen tutkimukseen ja siihen perustuvaan argumentointiin liittyviä ongelmakohtia. Valtaosan näistä ongelmakohdista tulisi olla kaikille tilastotiedettä harrastaville jotakuinkin selviöitä (esim. tilastollisen merkitsevyyden sopimuksenvaraiset raja-arvot:  $p < 0.001$ ,  $p < 0.5$  tai  $p < 0.1$ ). Kuitenkin teos myös osoittaa, että jopa hyvissä julkaisusarjoissa julkaistaan usein väärinymmärrettyjä huomioita ja paljon tilastollisessa mielessä heikkokkoa tutkimusta. Ketokiven vertaisilleen esittämä kritiikki on osittain kärkeästä, mutta samalla se on myös hyvin perusteltua.

Pelkkä tutkijoiden tilastomatemaattisen asiantuntemuksen arvostelu ja sittäminen ei kuitenkaan ratkaise yhtään ongelmaa. Ketokivi käyttääkin huomattavan osan kirjastaan näiden ongelmien tarkempaan valoittamiseen hyödyllisten ”parannusehdotuksien” ja esimerkillisesti toteutettujen tutkimusten kautta. Nämä esimerkillisesti toteutetut tutkimukset ovat monesti Ketokiven itsensä (joko yksinään tai yhteistyössä muiden suomalaisten ja kansainvälisten tutkijoiden kanssa) julkaisemia. Tämä on erityisen tärkeää. Muuten Ketokiven näkökulma lipsahtaisi pelkäksi hedelmättömäksi kritiikiksi, jossa muiden tutkijoiden tulokset altistetaan (hänen itsensä kuvaileman) skeptikon kiusattavaksi. Tällainen skeptikko ei anna tutkijalle mahdollisuutta olla oikeassa, koska hänen tutkimukselle asettamansa korkeat kriteerit tekevät tämän mahdottomaksi.

Joissain paikoin teosta kuitenkin vaivaa mielestäni tietoinen huonoimpien mahdollisten esimerkkien valitseminen. Lukija muun muassa tutustutetaan innovaatioiden mittaamisen heikkouksiin tutkimus- ja kehittämistoimintaan (t&k) käytettyjen panostusten, tai tarkemmin t&k-intensiivisyyden (esimerkiksi yrityksen tai valtion t&k-budjetti suhteessa yrityksen liikevaihtoon tai valtion bruttokansantuotteeseen) kautta. Ketokivi listaa ansiokkaasti ja aivan oikein t&k-mittaristojen heikkouksia päätyn lopputulemaan: t&k-intensiivisyys kuvaa heikosti yritysten (tai valtioiden ja alueiden) ”aitoa” innovatiivisuutta. Olen kirjoittajan kanssa asiasta täysin samaa mieltä. Kuitenkin innovatiivisuuden tai innovaatioihin liittyvässä nykykirjallisuudessa luotetaan enää kovin harvoin pelkästään t&k-intensiivisyyden mittaamiseen. Ennenminkin t&k-panostuksiin liittyviä tunnuslukuja käytetään yhtenä mittarina muiden nykyisesti yleisemmin käytössä olevien mittareiden (esim. patenttien) ohessa. Innovatiivisuus ei ole lineaarinen prosessi, jolla olisi

suoranaisesti alku- tai loppupistettä vaan se on jatkuvaa kehitystä eteen- ja taaksepäin suuntautuvine kytkeäntöineen. Kun innovatiivisuutta ajatellaan prosessina, ovat myös t&k-panostukset tärkeä, tosin epätyydellinen mittaristo, joka auttaa luomaan kokonaiskuvaa ilmiöstä. Ketokiven näkemys t&k-intensiivisyyden heikkouksista innovatiivisuuden mittarina on täten mielestäni hienaan tarkoitushakuisesti yksinkertaistettu.

Lisäksi teosten eri osioiden välillä tuntuu ajoittain vallitsevan pieniä ristiriitaisuuksia. Yhtäällä Ketokivi toteaa, että tieteen edistyminen vaatii uusien ongelmien ratkomista, mutta taas toisaalla hän pitää replikaatiotutkimuksen osuuden pienuutta kaikesta tutkimuksesta ongelmallisena tutkimustulosten yleistämiseksi. Samoin teoksessa harmitellaan aineistojen siivoamista ennen analyysin tekoa (esimerkiksi poistamalla ääriarvoja normaalijakauman saavuttamiseksi), koska samalla aineistosta saatetaan ajattelemattomasti poistaa

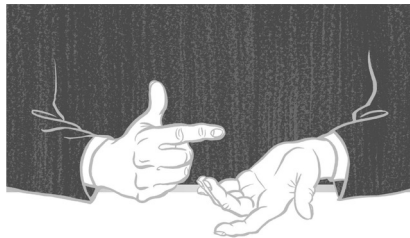
kaikkein mielenkiintoisimmat tapaukset. Toisaalla kuitenkin kehoitetaan ”trimmaamaan” aineistoa ennen kuvioiden piirtämistä. Havaitsemiani ”heikkouksia” ei tosin voi pitää kovinkaan merkittävinä ajatellen kirjan tarkoitusta ja yleis-sanomaa.

Ketokiven teoksen pohjalta voidaan lopuksi todeta, että tilastollinen malli on parhaimmillaan vain pintaraapaisu todellisen maailman monimutkaisista ja moniselitteisistä ilmiöistä; ei täydellinen, mutta hyvin toteutettuna tyhjä parempi. Lohdullisesti teos myös myöntää, että sen empiirille tutkimukselle ja tilastolliselle päättelylle asettamat standardit on tarkoituksellisesti asetettu mahdollisimman korkealle. Yksikään tutkija ei pysty ratkaisemaan kaikkia mahdollisia empiirisen tutkimuksen ongelmia ja tekemään tutkimuksestaan täysin aukotonta. Ketokivi ei ole tässä suhteessa idea-

listi, vaan hänen ajattelunsa perustuu ennenminkin teemoihin, joiden tulisi olla tiedemaailmassa yleisesti hyväksytyjä. Jokaisessa tutkimuksessa on hyvä pyrkiä täydellisyyteen, niin mahdotonta kuin sen saavuttaminen onkin, koska viime kädessä kaikki vastuu ja kunnia yksittäisen argumentin esittämisestä kuuluu tutkijalle. Itse tiede on kuitenkin luonteeltaan kumulatiivista, eikä vastuu sen edistämisestä näin ollen lepää yksinomaan yksittäisen tutkijan harteilla.

TEEMU MAKKONEN

*School of Hospitality and Tourism Management,  
University of Surrey &  
Department of Border Region Studies,  
University of Southern Denmark*



Mikko Ketokivi

## TILASTOLLINEN PÄÄTTELY JA TIETEELLINEN ARGUMENTOINTI

