

Mikko Ketokivi

# TILASTOLLINEN PÄÄTTELY JA TIETEELLINEN ARGUMENTOINTI



GAUDEAMUS

*Isälleni Heikille, joka opetti minulle mittaamisen taidon.  
Äidilleni Riitalle, joka opetti minut kysymään miksi.*

Copyright © 2015 Tekijä ja Gaudeamus Helsinki University Press

Kansi: Hanna Lahti / Huomen GDI

Gaudeamus Oy  
HYY Yhtymä  
[www.gaudeamus.fi](http://www.gaudeamus.fi)

KL 31.01  
UDK 311

ISBN 978-952-495-350-4

Painopaikka: Printon Trükikoda, Tallinna 2015

# SISÄLLYS

<b>JOHDANTO</b>	17
<b>1. TILASTOLLINEN PÄÄTTELY JA TEOREETTINEN SELITTÄMINEN</b>	31
1.1 Kuusi mekanismia	32
1.2 Päättele ei ole selittämistä	42
1.3 Tulkinnan haaste	46
1.4 Esimerkki: toimitusjohtajan sukupuoli ja yrityksen taloudellinen tulos	49
1.5 Argumentin rakenne: mitä, miten, miksi ja missä?	64
1.6 Argumentoinnin kollektiivinen ulottuvuus	67
1.7 Tilastollinen päättele osana tieteellistä tutkimusta	70
1.8 Yhteenveto: tilastollisen päätteilyn kriteerit tieteellisessä argumentissa	79
<b>2. OPERATIONALISOINTI JA MITTAAMINEN</b>	80
2.1 Teoreettiset ja empiiriset käsitteet	80
2.2 Korrespondenssisäännöt: operatiivisesta määritelmästä tosiarvoteoreettiseen tulkintaan	86
2.3 Hypoteesit tilastollisina assosiaatioina	89
2.4 Mittauksen luotettavuus	96
2.5 Reliabiliteetti	98
2.6 Validiteetti	104
2.7 Dimensionaalisuus ja konvergenssivaliditeetti: esimerkki älykkyystutkimuksesta	119
2.8 Määrittävät (formatiiviset) indikaattorit	121
2.9 Yleistettävyysteoria	126
2.10 Yhteenveto	129
<b>3. VARIANSSIN SELITTÄMINEN</b>	131
3.1 Varianssin selittäminen	132
3.2 Regressioanalyysin eri muodot	134
3.3 Tilastollinen päättele regressiofunktioista	137
3.4 Regressioanalyysin oletukset ja niiden merkitys käytännön tutkimuksessa	142
3.5 Mallin selitysaste $R^2$	149
3.6 Regressioanalyysin virhetermiin liittyvät oletukset	151
3.7 Piilomuuttujaregressio	154
3.8 Kontrollimuuttujat	158
3.9 Regressioanalyysi kategorisella selitettävällä muuttujalla	168
3.10 Yhteenveto	172

<b>4. KOVARIANSSIRAKENTEEN SELITTÄMINEN</b>	174
4.1 Kovarianssirakenne empiirisenä tutkimuskohteena	177
4.2 Parametritason näkökulma	180
4.3 Mallitason näkökulma	183
4.4 Mallin estimointi	186
4.5 Rakenneyhtälömalli faktorianalyysin ja polkuanalyysin yhdistelmänä	194
4.6 Yhteenveto	199
<b>5. SEURANTA-AINEISTOJEN ANALYYSI</b>	200
5.1 Prosessiselitykset ja varianssiselitykset	203
5.2 Autoregressiivinen tutkimusote	206
5.3 Yksilötason muutosmallit	210
5.4 Lineaarisen ja epälineaarisen latentin muutosmallin tarkastelu rakenneyhtälömallina	216
5.5 Interventiomalli latenttina muutosmallina	221
5.6 Muutosanalyysi luokkamuuttujilla: Markovin ketjut	226
5.7 Yhteenveto	230
<b>6. TULOSTEN TULKINTA</b>	232
6.1 Tulkinnan viitekehys	233
6.2 Yleistäminen	241
6.3 Vaihtoehtoiset selitykset ja tulkinnan kritiikki	243
6.4 Argumentin hyväksyminen tieteelliseksi tiedoksi	248
6.5 Yhteenveto	253
<b>7. TILASTOLLISEN TUTKIMUKSEN AVOIMET KYSYMYKSET</b>	254
7.1 Haaste 1: Vaihtoehtoisten selitysten ongelma	254
7.2 Haaste 2: Endogeenisuuden ongelma	257
7.3 Haaste 3: Yksimielisyyden illuusio	263
7.4 Haaste 4: Nyrkkisäännöistä luopuminen	265
7.5 Haaste 5: Korrespondenssien virheellinen mallinnus	268
7.6 Haaste 6: Oikean tilastollisen työkalun valinta	269
7.7 Haaste 7: Käyttäjystävälliset tilasto-ohjelmistot	274
7.8 Yhteenveto	277
7.9 Loppusanat	285
Liite A. Sanasto ja termien englanninkieliset käännökset	288
Liite B. Tilastollisten mallien graafinen esitystapa	296
Liite C. Identifioituvuus ja estimointi	300
Kirjallisuus	320
Kirjoittaja	336