

## Tilastomatemattiset menetelmät tutkimuksen työkaluina



Tilastotieteen perusmenetelmät sisältyvät nykyään lähes kaikkiin tutkijoiden käyttämiin laskentaohjelmistoihin.

Näennäisesti helppokäyttöisillä, graafisilla käyttöliittymillä varustetuilla tilasto-ohjelmistoilla saa tehtyä mitä monimutkaisimpia analyyseja helposti, nopeasti ja vaivattomasti. Tilasto-ohjelmisto ei kuitenkaan ole tilastotieteellinen menetelmä itse eikä ohjelmiston käyttö sinänsä tarkoita tilastomatemattisen menetelmän hallintaa. Tilanne on vastaava kuin, jos oletetaan, että pelkkä tekstinkäsittelyohjelmisto olisi riittävä hyvän tekstin tuottamiseen. Ei tekstinkäsittelyohjelmasta ole hyötyä, jos ei osaa kirjoittaa! Tilastotieteellisten menetelmien ymmärtäminen on ensiarvoisen tärkeää oikean analyysimenetelmän valinnassa sekä tilastotieteellisen päättelyn tekemisessä ja tulosten tulkitsemisessä.

Tilastotieteelliset mallit ja menetelmät pohjautuvat moniin eri oletuksiin. Vaikapa tavanomaisessa regressiomallinnuksessa saatetaan olettaa jäännösten olevan riippumattomia ja niille saatetaan olettaa vakio varianssi. Lisäksi saatetaan olettaa jäännösten olevan normaalisti jakautuneita. Tehtyjen oletusten paikkansapitävyyden analysointi ja raportointi ovat aina tutkijan vastuulla. Tilastotieteellisessä mallinnuksessa joudutaan usein tilanteeseen, jossa malli on sovitettava ennen kuin oletusten paikkansapitävyyttä voidaan tarkastella. Tämä johtaa siihen, että ikään kuin hakuammunnalla sovitetaan erilaisia malleja, ja niistä valitaan se, jossa oletukset näyttäisivät parhaiten toteutuvan. Tällöin oletusten paikkansapitävyyden tarkastelu saattaa unohtua ja sen merkitys hämärtyä, koska malli on jo estimoitu ja tutkijalla on kiire tulkitsemaan tuloksia.

Oletusten paikkansapitävyyden tarkastelemiseksi on tiedettävä mitä olettaa. Ymmärrys siitä miten oletukset liittyvät käytettyyn testiin tai menetelmään antaa lisäksi edellytyksiä arvioida, kuinka vakavia ja minkälaisia seurauksia oletusten paikkansapitämättömyydellä voi olla tuloksiin. On usein mahdollista arvioida, onko saatu testin merkitsevyysarvo yli- vai aliarvio. Oletusten paikkansapitävyyden merkitykseen vaikuttaa myös se, laaditaanko malli hypoteesin testausta, ilmiön selittämistä vai ennustamista varten. Menetelmän tuntemus antaa myös mahdollisuuksia vaihtoehtoihin. Regressioanalyysitilanteessa voidaan harkita, lähdetäänkö jännösvarianssin vakioimiseksi tekemään muunnoksia vastemuuttujaan ja luovutaan soviteen harhattomuudesta, vai tehdäänkö oletusarvoa todenmukaisempia oletuksia jännösvarianssista.

Tilastomatemaattiset menetelmät ovat nykyisin monen tutkijan työkaluja. Niiden käyttö ei kuitenkaan saa olla itsetarkoitus. Hienoimmillakaan tilastotieteellisillä menetelmillä ei voida korvata tutkijan luovaa ajattelua ja terveen järjen käyttöä. Hyvän tutkimuksen tunnusmerkki ei ole monimutkaisten menetelmien käyttö, vaan uutta luova ajattelu, järkevien hypoteesien ja teorioiden luominen sekä niiden paikkansapitävyyden testaaminen tehokkailla menetelmillä. Tilastotieteelliset menetelmät ovat usein päteviä ja tehokkaita analyysimenetelmiä, jotka oikeuttavat tilastotieteellisen päättelyn avulla tuloksen yleistettävyyden. Siksi tilastotieteellisten menetelmien teorian ymmärtämisellä ja menetelmien hallinnalla on oleellinen merkitys tutkimuksessa. Tutkijan tehtäväksi kuitenkin aina jää tilastollisesti merkitsevän tuloksen tai vaikutuksen merkittävyyden tulkitseminen ja johtopäätösten tekeminen reaali maailman ilmiöstä.

*Lauri Mehtätalo*

MTT Lauri Mehtätalo on Metsätieteen aikakauskirjan avustava tieteellinen toimittaja. Hän on Helsingin yliopiston metsäbiometrian dosentti.