Kätevä arvio riskisuhteesta avuksi yhteiskuntatieteilijöille

PEKKA PERE

Riski- ja ristisuhde sekoitetaan usein. Riskisuhteen suuruusluokka voidaan tietyin edellytyksin arvioida helposti ristisuhteesta. Kaava voi auttaa kvantitatiivisten yhteiskuntatieteellisten tutkimusten lukijoita. Eniten lukijoita auttaisi muutos tulosten raportointikäytännössä.

Vääriä tulkintoja

Hannu Karhunen ja Roope Uusitalo (2017) kiinnittivät tässä lehdessä huomiota tilastotieteellisen *odds ratio* -käsitteen kommervenkkiuteen. He arvioivat, että sen ”tulkinta on useimmille lukijoille – ja ajoittain myös kirjoittajille – vaikeaa” ja katsoivat, että ”todennäköisyyksien välinen suhde eli ns. riskisuhde olisikin useimmille paljon helpommin ymmärrettävä luku”. Kirjoittajat viittasivat esimerkkeihin riskisuhteen (*risk ratio* tai *relative risk*) ja *odds ratio*n sekoittamisista.

Kutsun *odds ratio*ta ristisuhteeksi (piakkoin julkaistavan Suomen Tilastoseuran Tilastotieteen sanaston suomennos). Esitän alla kaksi konkreettista esimerkkiä riski- ja ristisuhteen sekoittamisesta.

Yhteiskuntatieteilijän haastattelussa (Acatiimi 6/2018) sekoitettiin riskisuhde

$$\lambda = \frac{\pi_1}{\pi_2}$$

ja ristisuhde

$$\theta = \frac{\frac{\pi_1}{1-\pi_1}}{\frac{\pi_2}{1-\pi_2}} = \frac{\pi_1}{\pi_2} \times \frac{1-\pi_2}{1-\pi_1} = \lambda \times \frac{1-\pi_2}{1-\pi_1}. \quad (1)$$

Yllä π_1 ja $1-\pi_1$ ovat tapahtuman ja sille vastaisen tapahtuman todennäköisyydet (esim. yhteiskuntaluokkaan 1 kuuluva pääsee opiskelemaan tai ei pääse). π_2 ja $1-\pi_2$ ovat vastaavat todennäköisyydet toiselle tapahtumalle (yhteiskuntaluokkaan 2 kuuluva pääsee opiskelemaan tai ei pääse).

Haastattelussa kirjoitettiin: ”Tutkimusten mukaan akateemisten perheiden lapset päätyvät seitsemän kertaa useammin yliopistoon opiskelemaan kuin ei-akateemisten lapset.” Ajatus perustui tutkimukseen, jonka mukaan akateemisten ja ei-akateemisten perheiden jälkeläisten yliopisto-opintoihin osallistumisen eroja mittaavan ristisuhteen estimaatti on 6.8 (noin 7). (Kivinen ym. 2012.) Ristisuhteen θ ollessa 7 ylipäänsä λ ei ole 7 eikä todennäköisyys π_1 ole seitsenkertainen π_2:een verrattuna. Aiheesta on uutisoitu vastaavaa väärin tulkittua tietoa myös aiemmin (Aamulehti ja Kaleva 5.3.2016).

Sosiaalityön väitöskirjan taulukossa 21 raportoitiin ristisuhteen estimaatit 4.5, 4.2, 2.1 ja 4.2 (Väänänen 2013). Väittelijä tulkitsi (mts. 93): ”Yhteenvetona voidaan todeta, että mikäli lapsen perhe ei ole lapsen kahden biologisen vanhemman perhe, riski psyykkisille ongelmille on 4.5- kertainen [sic]. Mikäli perhedynamiikan toimivuus on keskiarvoa alhaisempi, riski psyykkisille häiriöille on 4.2- kertainen ja mikäli sosiaaliset ja emotionaaliset arvot perheessä ovat keskimääräistä alhaisemmat, riski häiriöille on kaksinkertainen. Lisäksi pojilla on tyttöihin verrattuna 4.2- kertainen riski psyykkisille häiriöille.” Väitöskirjassa samaistettiin riski- ja ristisuhde samalla tavalla virheellisesti kuin osoittamassani haastattelussa.

Sekaannus voi olla vaarallinen, koska riski- ja ristisuhde voivat erota suuresti ja tutkitut asiat ovat yhteiskunnallisesti tärkeitä. Sekaannus on yleinen tilastotieteen soveltajien keskuudessa eri tieteenaloilla Suomessa ja kansainvälisesti (esim. Feng ym. 2019; Pere ym. 2017; Rita & Virtala 2013).

Ominaisuuksia

Ristisuhdenimitys juontaa nelikentästä, jonka kenttien todennäköisyydet kerrotaan kulmittain ristiin ja suhteutetaan:

π_1	$1-\pi_1$
π_2	$1-\pi_2$

($[\pi_1(1-\pi_2)]/[\pi_2(1-\pi_1)] = [\pi_1/(1-\pi_1)]/[\pi_2/(1-\pi_2)] = \theta$). Rita ym. (2008) ovat selittäneet risti- ja riskisuhteen eron huolella. Feng ym. (2019) avaavat niiden yhteyksiä ja eroja yksityiskohtaisesti. Tarkastelut alla havainnollistavat.

Riskisuhde voi saada saman arvon monilla π_1:n ja π_2:n yhdistelmillä. Esimerkiksi parit $\pi_1 = 0.07$ ja $\pi_2 = 0.01$, $\pi_1 = 0.49$ ja $\pi_2 = 0.07$ sekä $\pi_1 = 0.84$ ja $\pi_2 = 0.12$ tuottavat arvon $\lambda = 7$.

Edellä annetuilla todennäköisyyksien π_1 ja π_2 arvoilla $= 7$ mutta $\theta = 7.5$, $\theta = 12.8$ tai $\theta = 38.5$.[1] Riski- ja ristisuhde voivat saada täysin toisistaan poikkeavia arvoja.

Jos $\pi_1 > \pi_2$, niin riskisuhde on suurempi kuin 1 ja ristisuhde on suurempi kuin riskisuhde:

$$1 < \lambda < \theta \, .$$

Oikeanpuoleisen epäyhtälön sanoma on yksinkertainen: Jos ristisuhde sekoitetaan riskisuhteeseen, riskisuhde arvioidaan liian suureksi ($\pi_1 > \pi_2$ -tilanteessa). Esimerkkiluvut toteuttavat epäyhtälöt ($1 < 7 < 7.5$, $1 < 7 < 12.8$ tai $1 < 7 < 38.5$) ja osoittavat, että ristisuhteen suuretessa riskisuhde ei välttämättä suurene. Helppo muistisääntö tieteellisten artikkelien lukijoille on, että ristisuhde on yläraja riskisuhteelle, jos $\pi_1 > \pi_2$.

Ristisuhteen oikeanpuolimmaisimmasta muotoilusta kaavassa (1) nähdään, että risti- ja riskisuhde voivat olla lähes samoja, jos todennäköisyydet ovat hyvin pieniä. Tällöin osamäärä $(1 - \pi_2)/(1 - \pi_1)$ on noin 1 ja θ on noin λ. Riski- ja ristisuhde yhtyvät, jos $\pi_1 = \pi_2$. Tällöin $\lambda = \theta = 1$. Ylipäänsä riski- ja ristisuhde eroavat.

Kätevä arvio ja suurin apu

Tyler VanderWeele (2017) esittää helpon tavan arvioida riskisuhdetta ristisuhteen avulla. Oletetaan selkeyden vuoksi, että $\pi_1 > \pi_2$.

VanderWeelen oivallus on pohtia tilannetta, jossa todennäköisyydet ovat yhtä lähellä 0.5:ttä: $\pi_1 = 0.5 + w$ ja $\pi_2 = 0.5 - w$, joissa $0 < w < 0.5$. Tällöin risti- ja riskisuhdetta sitoo yhtäsuuruus $\theta = \lambda^2$:

$$\theta = \frac{(0.5+w)}{(0.5-w)} \times \frac{1-(0.5-w)}{1-(0.5+w)} = \frac{(0.5+w)^2}{(0.5-w)^2} = \lambda^2.$$

Riskisuhde selviää nyt tarkasti ja helposti laskemalla neliöjuuri ristisuhteesta: $\lambda = \sqrt{\theta}$. Esimerkki: Jos $w = 0.3$, niin $\pi_1 = 0.8$, $\pi_2 = 0.2$, $\theta = 16$ ja $\lambda = \sqrt{16} = 4$. VanderWeele ehdottaa, että soveltajat voisivat käyttää $\sqrt{\theta}$:a karkeana arviona λ:sta yleisemminkin tilanteissa, joissa molemmat todennäköisyydet π_1 ja π_2 ovat 0.2:hta suurempia. Näin laskettu arvio riskisuhteesta olisi yleensä parempi arvio siitä kuin riskisuhteeksi kuviteltu ristisuhde. Esimerkki: Jos $\pi_1 = 0.6$ ja $\pi_2 = 0.3$, niin $\lambda = 2$, $\theta = 3.5$ ja $\sqrt{3.5} = 1.9$. Neliöjuurilaskun tuottama 1.9 on lähempänä oikeaa riskisuhdetta 2 kuin riskisuhteeksi väärintulkittu 3.5.

Yliopisto-opiskeluesimerkin ristisuhteen 6.8 neliöjuuri on 2.6. Perhemuotoesimerkin ristisuhteen 4.5 neliöjuuri on 2.1. Riskisuhteiden neliöjuuriarviot eroavat huikeasti ristisuhteista. Jos yhteiskuntapolitiikkaa tehtäisiin väärinkäsitettyjen ristisuhteiden perusteella, toimet voisivat olla suuresti ylimitoitettuja tai täysin perusteettomia.

Lukuihin 2.6 ja 2.1 ei pidä suhtautua vakavasti otettavina empiirisinä riskisuhdearvioina. Neliöjuurikaavan oletus riittävän suurista todennäköisyyksistä ei välttämättä päde näissä esimerkeissä. Havainnollistan vain, kuinka vakava virheellinen tulkinta on mahdollinen, jos ristisuhde tulkitaan riskisuhteeksi.

Neliöjuurikaava $\lambda = \sqrt{\theta}$ voi auttaa tutkimusta luettaessa. Sillä saa kätevästi näkemyksen riskisuhteen suuruusluokasta, jos tutkimuksessa on raportoitu vain ristisuhde ja jos todennäköisyydet taustalla ovat riittävän suuret.

Tutkija voi laskea todennäköisyyksiä ja riskisuhteita estimoimastaan mallista, vaikka tilasto-ohjelmisto raportoisi ristisuhteen muttei todennäköisyyksiä. Informatiivisinta ja suositeltavinta olisi, että tutkija raportoisi ristisuhteen lisäksi sen taustalla olevia todennäköisyyksiä ja riskisuhteita. Ne hahmottuvat ristisuhdetta helpommin ja olisivat suurin apu lukijoille.

1 Esitän laskujen lopputulokset desimaalin tarkkuudella.

KIRJALLISUUS

Aamulehti (2016) Akateemisista kodeista lääkäriksi yli 10 kertaa useammin – Näin koulutus periytyy eri aloilla. 5.3.2016.

Acatiimi (2018) Hyvinvointiyhteiskunnan todellisuuden tutkija. 6/2018, 24–29.

Feng, Changyong & Bokai Wang & Hongyue Wan (2019) The relations among three popular indices of risks. Statistics in Medicine, 38, 23, 4772–4787.

Kaleva (2016) Akateemisista kodeista lääkäriksi yli 10 kertaa useammin — Näin koulutus periytyy eri aloilla. 5.3.2016.

Karhunen, Hannu & Roope Uusitalo (2017) 50 vuotta koulutusmahdollisuuksien tasa-arvoa. Yhteiskuntapolitiikka 82, 296–303.

Kivinen, Osmo & Juha Hedman & Päivi Kaipainen (2012) Koulutusmahdollisuuksien yhdenvertaisuus Suomessa. Eriarvoisuuden uudet ja vanhat muodot. Yhteiskuntapolitiikka, 77, 5, 559–566.

Pere, Pekka & Eero Lilja & Anton Sobolev (2017) Maahanmuuttajien mielikuva naisten suosimisesta tuomioistuimissa. Lakimies, 115, 1, 90–99.

Rita, Hannu & Pertti Töttö & Marja Alastalo (2008) Voiko turkulaisten kirjoittamista artikkeleista yli 100 % olla kvantitatiivisia? Vetosuhteen (odds ratio) ja vedon (odds) tulkintaa. Janus, 16, 1, 72–80.

Rita, Hannu & Anna-Maija Virtala (2013) Termien odds ja odds ratio tulkinnasta: ”Veren alhainen immunoglobuliini ja G-pitoisuus kaksinkertaistaa vasikoiden keuhkokuumepaineen”. Suomen eläinlääkärilehti, 119, 2, 67–75.

VanderWeele, Tyler J. (2017) On a square-root transformation of the odds ratio for a common outcome. Epidemiology, 28, 6, e58–e60.

Väänänen, Riitta (2013) Perheen rakenteen, dynamiikan ja arvojen merkitys lapsen psyykkiselle hyvinvoinnille. Väitöskirja. Itä-Suomen yliopisto. Publications of the University of Eastern Finland, Dissertations in Social Sciences and Business Studies, 68.