Sokkotestit ja epäsuora kausaalinen päättely

Tommi Mononen

Helsinki 16.4.2003

Kausaalimallien tutkimusseminaari

HELSINGIN YLIOPISTO
Tietojenkäsittelytieteen laitos

1 Johdanto

Lääketieteellisissä kliinisissä testeissä käytetään usein niin sanottua sokkotestiä uuden lääkkeen vaikutuksen testaamiseen. Testissä vain osalle ihmisistä annetaan testattavaa lääkettä. Näiden ihmisten joukkoa kutsutaan testiryhmäksi. Lopuille ihmisistä, joiden muodostamaa joukkoa kutsutaan kontrolliryhmäksi, annetaan lumelääkettä. Kummankaan ryhmän jäsenet eivät tiedä kumpaan ryhmään he kuuluvat. Periaatteessa tällainen testausmenetelmä on tarkka ja luotettava vain, jos kaikki kokeeseen osallistuvat henkilöt noudattavat heille annettuja ohjeita lääkkeen käytöstä. Näin ei kuitenkaan tapahdu suurimmassa osassa niistä tapauksista, joissa henkilöitä ei suljeta tarkasti kontrolloituun ympäristöön. Parempi menetelmä onkin suositella lääkettä ja antaa ihmisten itse päättää sen käytöstä, jolloin kerätty aineisto on myös luotettavampaa [Pea00]. Itse testiasetelmasta voidaan muodostaa kausaalimalli [Pea95]. Mallin yhteensopivuudesta testiasetelman kanssa voidaan varmistautua jossain määrin tarkastelemalla mallin muuttujien välisiä suhteita. Kyseisen mallin avulla, ottamalla huomioon koehenkilöiden käyttäytyminen, voidaan laskea rajat lääkkeen positiiviselle tai negatiiviselle vaikutukselle. Toisaalta jos pystytään tekemään vahvempia oletuksia koehenkilöiden käyttäytymisestä, niin myös varmuus lääkkeen todellisesta vaikutuksesta kasvaa. Rajojen laskemiseksi joudutaan kuitenkin olettamaan havaitun testistä saadun datan todennäköisyysjakaumien noudattavan todellista jakaumaa. Pienillä testiaineistoilla kyseinen oletus on epärealistinen.

Kausaalimallista voidaan rakentaa Bayes-verkko lääkkeen vaikutuksen selvittämiseksi. Gibbsin otanta mahdollistaa lääkkeen vaikutusta kuvaavan jakauman laskemisen Bayes-verkolla. Bayes-mallin käyttämä priori vaikuttaa varsinkin pienillä otoskoilla vahvasti Gibbsin otannan tuottamaan posteriorijakaumaan. Annettua kausaalimallia ja edellä mainittuja estimaatteja lääkkeen vaikutuksesta voidaan myös käyttää kontrafaktuaalisten todennäköisyyksien selvittämiseen tietyn yksilön tapauksessa [ChP99].

2 Testiasetelma

Tarkastellaan ensiksi tavallisen sokkotestauksen suurimpia puutteita. Tämän jälkeen esitellään testauksen taustalla esiintyvä todellinen kausaalirakenne ja siihen liittyvät oletukset. Lopuksi tarkastellaan, kuinka kausaalirakenteessa esiintyvä latentti muuttuja pystytään esittämään äärellisenä muuttujana.

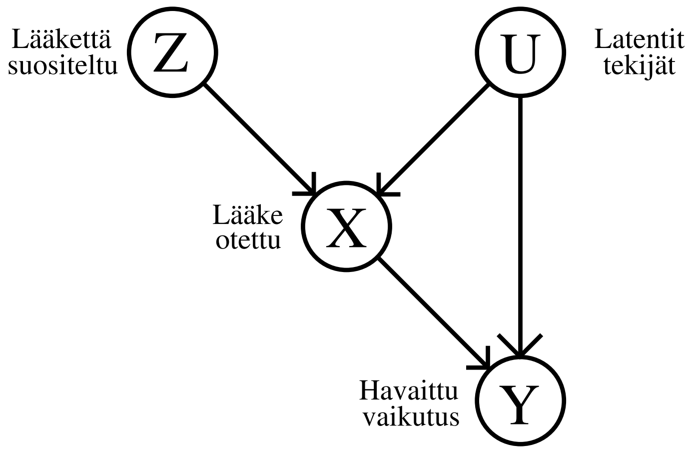
2.1 Sokkotestit

Tavallisen sokkotestin tuloksen oletetaan usein kuvaavan lääkkeen todellista vaikutusta. Näin ei kuitenkaan ole, sillä testiyksilöt saattavat jättää noudattamatta ohjeita annetun lääkkeen tai lumelääkkeen käytöstä [Pea95]. Osa testiryhmän jäsenistä saattaa salaa esimerkiksi lopettaa lääkkeen käytön, jos se aiheuttaa epämiellyttäviä sivuvaikutuksia. Jos nyt lääkkeen käytön lopettaneet henkilöt paranevat itsestään tai toisen salaa syödyn lääkkeen avulla, tulkitaan lääke todellista hyödyllisemmäksi. Sokkotesteihin osallistuvat henkilöt saattavat olla jotenkin erikoisesti valikoituneet tai heillä saattaa olla eri kannustimet käyttää lääkettä kuin normaalilla henkilöllä. Näissä tapauksissa voidaan olettaa lääkkeen todellisten hoitotulosten eroavan testin tuloksista.

Toinen luonnollisempi menetelmä testaamiseen on antaa satunnaisesti valituille kokeeseen osallistuville henkilöille valinnanvapaus. He saavat itse päättää suosittelun jälkeen käyttävätkö he lääkettä vai eivät. Menettely koskee siis sekä kontrolli- että testiryhmää. Henkilöt voivat myös itse päättää annosmäärän muuttamisesta tietyissä rajoissa. Lääkkeen käyttöä voidaan tällöin seurata esimerkiksi käyttämällä kulutusta mittarina. Tällaisen testin tapauksessa tietenkin lääkkeen todellisen vaikutuksen laskeminen on monimutkaisempaa. Toisaalta tavalliset sokkotestitkin voidaan luokitella kuuluvaksi tähän ryhmään, jos kaikki koehenkilöt eivät kuitenkaan noudata ohjeita lääkkeen käytöstä. Tavallisten sokkotestien tapauksessa tutkijat eivät kuitenkaan saa tietää lääkkeen käyttämättä jättämisestä. Koehenkilöillä tässä muutetussa testimenettelyssä ei kuitenkaan ole tarvetta valehdella lääkkeen käyttöä, mikä parantaa kokeen informatiivisuutta [Pea00].

2.2 Kausaalirakenne

Edellä esitelty muunnettu sokkokoe voidaan esittää kuvassa 1 näkyvällä kausaalirakenteella [Pea00]. Verkon solmut X,Y,Z ovat binäärisiä muuttujia ja solmu U kuvaa kaikkia muita kokeeseen vaikuttavia havaittuja tai havaitsemattomia tekijöitä. Havaittu muuttuja Z kuvaa satunnaistettua lääkkeensuosittelumenettelyä. Muuttuja X kertoo ottiko koehenkilö lääkettä vai ei ja muuttuja Y paraniko hän. Latentit tekijät U vaikuttavat sekä lääkkeen ottoon että havaittuun vaikutukseen, mutta ovat riippumattomia satunnaistetusta menettelystä Z.



Kuva 1: Kausaalisuhteet muunnetun sokkotestin tapauksessa.

Muunnetussa sokkotestissä ollaan kiinnostuttu muuttujien X ja Y välisestä suhteesta, mutta koska todellista lääkkeen ottoa ei voida havainnoida suoraan, joudutaan turvautumaan epäsuoraan menettelyyn. Satunnaistetun menettelyn avulla vapaudutaan muuttujan U aiheuttamasta sekoittumisesta muuttujien X ja Y välillä ja pystytään päättelemään edellä mainittujen muuttujien välisiä suhteita. Tämä on seurausta siitä, että Z katkaisee muuttujien U ja X välisen riippuvuuden [Pea00].

Merkitään z_1, kun koehenkilölle on suositeltu lääkettä ja z_0, kun koehenkilölle ei ole suositeltu lääkettä. Vastaavasti merkitään x_1, kun koehenkilö on ottanut lääkettä ja

x_0, kun koehenkilö ei ole ottanut lääkettä. Lääkkeen positiivista vaikutusta merkitään arvolla y_1 ja negatiivista vaikutusta arvolla y_0. Merkitään lisäksi $x \in \{x_0, x_1\}$, $y \in \{y_0, y_1\}$ ja $z \in \{z_0, z_1\}$.

2.3 Latentin muuttujan arvojoukko

Latentti muuttuja U voidaan aina esittää laskennallisesti ekvivalenttina muuttujana, jolla on äärellinen määrä tiloja riippumatta tekijän U taustalla esiintyvästä monimutkaisuudesta. Aikaisemmin esitellyn mallin, jossa X, Y ja Z ovat binäärisiä, tapauksessa tilojen lukumäärä on 16. Neljä funktiota kuvaa muuttujien Z ja X välistä funktionaalista suhdetta. Samoin neljä funktiota kuvaa muuttujien X ja Y välistä suhdetta. Näiden suhteiden karteesisen tulon muodostamien alkioiden lukumäärä on tuo mainittu luku 16. Muuttujien Z ja X väliset funktiot voidaan tässä tapauksessa tulkita ihmistyypeiksi: kieltäytyjä ($x = 0$), vastaanhangoittelija ($x \neq z$), tottelevainen ($x = z$) ja aina lääkettä syövä ($x = 1$). Samalla tavoin voidaan kuvata lääkkeenoton ja paranemisen välinen yhteys. Saadaan potilastyypit: parantumaton ($y = 0$), lääkkeestä vahingoittuva ($y \neq x$), lääkkeen avulla parantuva ($y = x$) ja aina paraneva ($y = 1$). Kukin äärellisen muuttujan tila on siis ihmistyypin ja potilastyypin muodostama pari. Kun tiedetään latentin tekijän tilat, voidaan niitä käyttää apuna matemaattisessa analyysissä ja johtaa seuraavassa luvussa esitellyt eksaktit rajat lääkkeen vaikutukselle [Pea00].

3 Testisuureet ja niiden laskeminen

Luvussa esitellään keskeiset testisuureet ja niiden tuottamat teoreettiset rajat lääkkeen negatiiviselle tai positiiviselle vaikutukselle. Lisäksi keskitytään lääkkeen vaikutusta kuvaavan jakauman arviointiin.

3.1 Keskimääräinen kausaalivaikutus

Lääketesteissä mitataan lääkkeen käytön ja vaikutuksen välistä suhdetta. Ilmaistuna kausaalikäsitteiden avulla, pyritään siis mittaamaan lääkkeen käytön X ja havaitun vaikutuksen Y välinen *keskimääräinen kausaalinen vaikutus* (average causal effect). Tämä suure voidaan ilmaista lausekkeella

$$\begin{aligned} ACE(X \to Y) &= P(y_1|do(x_1)) - P(y_1|do(x_0)) \\ &= \sum_u [P(y_1|x_1, u) - P(y_1|x_0, u)]P(u), \end{aligned} \tag{1}$$

Missä u on muuttujan U tila. Lausekkeen arvo kuvaa parantumistodennäköisyyksien erotusta lääkkeen käytön ja käyttämättömyyden välillä. Kyseinen arvo on juuri sama mihin tavallisella sokkotestillä päästään ideaalitapauksessa [Pea00].

Keskimääräisen kausaalivaikutuksen testisuure ei kuitenkaan kelpaa selllaisenaan kaikissa tapauksissa. Jos oletetaan lääkkeen olevan yleisesti käytössä ja halutaan tietää pitäisikö lääkkeen määräämistä potilaille jatkaa, tulee tällöin laskea lääkkeen vaikutus hoidettujen ryhmässä. Testisuureeksi [Pea00] saadaan tällöin

$$ACE^*(X \to Y) = \textstyle\sum_u [P(y_1|x_1, u) - P(y_1|x_0, u)]P(u|x_1).$$

Testisuure kuvaa siis ehdollista odotusarvoa u:n suhteen ehdolla $X = x_1$.

3.2 Luonnolliset rajat

Niin sanotut luonnolliset rajat [Pea95] ovat yksinkertaisempi menetelmä lausekkeessa 1 määritellyn keskimääräisen kausaalisen vaikutuksen arviointiin kuin tarkat rajat. Luonnolliset rajat määritellään epäyhtälöiden

$$\begin{aligned} ACE(X \to Y) &\geq P(y_1|z_1) - P(y_1|z_0) - P(y_1, x_0|z_1) - P(y_0, x_1|z_0), \\ ACE(X \to Y) &\leq P(y_1|z_1) - P(y_1|z_0) + P(y_0, x_0|z_1) + P(y_1, x_1|z_0) \end{aligned} \tag{2}$$

avulla. Ensimmäiset kaksi termiä ovat kummassakin epäyhtälössä samat. Ne yhdessä kuvaavat lääkkeenottorohkaisun tai -kehoituksen vaikutusta paranemiseen.

Luonnolliset rajat eivät tuota yhtä kapeaa varmuusväliä kuin seuraavaksi esiteltävät tarkat rajat. Luonnollisten rajojen varmuusvälin leveys kuvaa vastaanhangoittelun määrää, joka on todennäköisyys $P(x_1|z_0) + P(x_0|z_1)$. Jos kuitenkin voidaan varmasti osoittaa, ettei yksikään koehenkilö valitse tahallaan päinvastoin kun on suositeltu, ovat luonnolliset rajat optimaaliset [Pea00]. Luonnollisen rajojen käyttöä tukee toisaalta myös niiden laskemisen yksinkertaisuus.

3.3 Tarkat rajat

Tarkat rajat [Pea00] saadaan määrättyä käyttämällä apuna edellä mainittua muuttujasta U muodostettua äärellistä esitystä. Kun keskimääräinen kausaalinen vaikutus lausutaan tämän äärellisen muuttujan tilojen avulla ja lopuksi käytetään lineaarista optimointia ottaen huomioon todennäköisyyksien summautumisen yhteen, saadaan alla esitellyt rajat. Rajojen esitysasun yksinkertaistamiseksi valitaan lyhennysmerkinnäksi $p_{ab.c} = P(y_a, x_b|z_c)$. Nyt voidaan esittää tarkat rajat matemaattisessa muodossa. Alarajaksi saadaan

$$ACE(X \rightarrow Y) \geq max \left\{ \begin{array}{l} p_{11.1} + p_{00.0} - 1 \\ p_{11.0} + p_{00.1} - 1 \\ p_{11.0} - p_{11.1} - p_{10.1} - p_{01.0} - p_{10.0} \\ p_{11.1} - p_{11.0} - p_{10.0} - p_{01.1} - p_{10.1} \\ -p_{01.1} - p_{10.1} \\ -p_{01.0} - p_{10.0} \\ p_{00.1} - p_{01.1} - p_{10.1} - p_{01.0} - p_{00.0} \\ p_{00.0} - p_{01.0} - p_{10.0} - p_{01.1} - p_{00.1} \end{array} \right\} \tag{3}$$

ja ylärajaksi saadaan

$$ACE(X \to Y) \leq min \left\{ \begin{array}{l} 1 - p_{01.1} - p_{10.0} \\ 1 - p_{01.0} - p_{10.1} \\ -p_{01.0} + p_{01.1} + p_{00.1} + p_{11.0} + p_{00.0} \\ -p_{01.1} + p_{11.1} + p_{00.1} + p_{01.0} + p_{00.0} \\ p_{11.1} + p_{00.1} \\ p_{11.0} + p_{00.0} \\ -p_{10.1} + p_{11.1} + p_{00.1} + p_{11.0} + p_{10.0} \\ -p_{10.0} + p_{11.0} + p_{00.0} + p_{11.1} + p_{10.1} \end{array} \right\}. \tag{4}$$

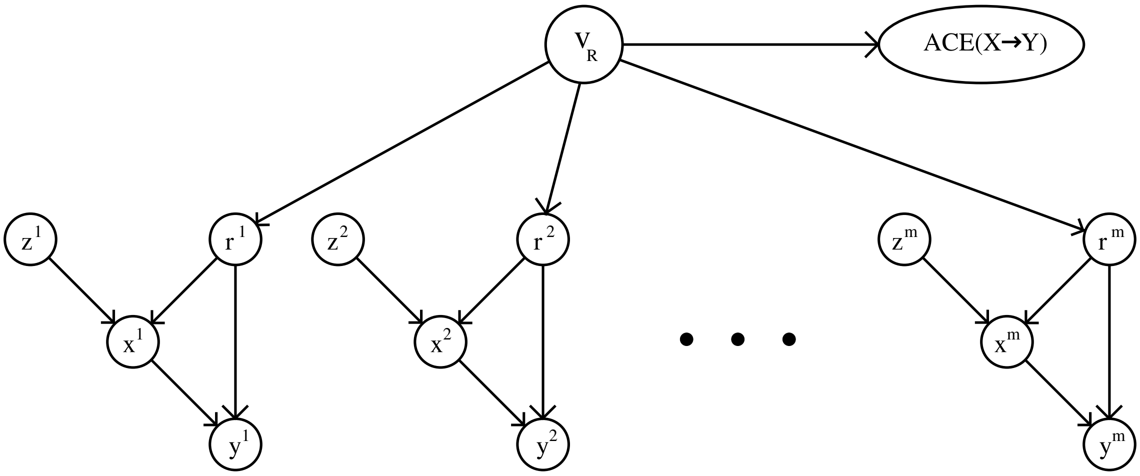
Näitä rajoja tiukempia rajoja ei voida saavuttaa ilman oletuksia eri potilas- ja ihmistyyppien muodostamien parien prosenttiosuuksista.

3.4 Lääkkeen vaikutuksen laskeminen bayesiläisittäin

Asymptoottisia rajoja käytettäessä valitaan usein se rajoista, joka antaa pessimistisemmän arvion. Jos ei haluta turvautua asymptoottisiin oletuksiin äärellisten otosten yhteydessä, voidaan laskea ratkaisu bayesiläisiä menetelmiä käyttäen. Tällöin saadaan ratkaisuksi keskimääräisen kausaalisen vaikutuksen jakauma. Mitä enemmän havaintoja on, sitä lähempänä tämä posteriorijakauma on oikeaa arvoa ja sitä keskittyneempi se on oikean välin sisälle tai oikean pisteen ympärille. Jos ratkaisu ei ole identifioituva, saadaan tulokseksi väli, vaikka näytteitä olisi rajaton lukumäärä. Priorijakauma joudutaan määrittelemään, mutta sen vaikutus häviää kokonaan, jos otosmäärä on suuri [ChP99].

Muunnetusta sokkotestiasetelmasta voidaan muodostaa kuvassa 2 näkyvä graafinen malli. Solmu $ACE(X \to Y)$ on kohdesolmu, jonka posteriorijakauma pyritään selvittämään. Kyseisen solmun ja solmun v_R suhde on deterministinen. Solmu v_R kuvaa ihmis- ja potilastyyppiparien todennäköisyyksiä ja siis näiden parien suhteita havaintoaineistossa. Solmut r^i, missä $i = 1..m$, esittävät kunkin potilaan tuntematonta ihmis- ja potilastyyppiä. Luku m on siis testissä olevien potilaiden lukumäärä.

Muiden mallin solmujen merkitys on sama kuin kuvan 1 solmujen.



Kuva 2: Gibbsin otannassa käytetty malli.

Ongelmasta muodostetussa bayesiläisessä mallissa on kuitenkin solmuja, joiden arvot eivät ole havaittuja. Solmujen r^i vuoksi päättelyalgoritmi ei voi laskea suoraan keskimääräisen kausaalisen vaikutuksen posteriorijakaumaa suoraan vaan joudutaan käyttämään Gibbsin otantaa. Tämän menetelmän avulla saadaan laskettua iteratiivisesti posteriorijakauma. Algoritmin tuottama posteriorijakauma on sitä tarkempi mitä lähempänä priorijakauma on todellista jakaumaa ja mitä enemmän aineistoa algoritmille annetaan. Tasaista jakaumaa on yleensä syytä käyttää priorijakaumana. Kun aineistoa on vähän, tulee mieluummin käyttää asymptoottisia rajoja päättelyyn, koska valitun priorin vaikutus on tällöin turhan suuri [ChP99].

4 Täydennyksiä

Lineaarisen optimoinnin avulla voidaan johtaa myös rajat kontrafaktuaalitapahtuman todennäköisyydelle. Voidaan siis kysyä, mitä olisi tapahtunut, jos tietty yksilö ei olisikaan ottanut lääkettä. Tämän jälkeen tarkastellaan vielä kausaalimallistruktuurin todenmukaisuuden testausta.

4.1 Kontrafaktuaalinen päättely

Kontrafaktuaalisessa päättelyssä [Pea00] ollaan kiinnostuttu selvittämään, että mikä olisi ollut seuraus, jos jotain asiaa ei olisikaan tapahtunut. Jotta voidaan puhua aidoista kontrafaktuaaleista, pitää olettaa ettei samaa alkutilannetta voida enää toistaa. Esimerkiksi kuollut ihminen ei voi enää jättää syömättä lääkettä, jonka syönti saattoi aiheuttaa juuri kyseisen ihmisen kuoleman.

Kun puhutaan kontrafaktuaalisesta päättelystä yksilön kohdalla, joudutaan turvautumaan vahvoihin oletuksiin, jotka todennäköisesti ovat vääriä [Daw97]. Todennäköisyyksien laskemiseksi tulee olettaa yksilön olevan tavallaan mitatun joukon keskiarvoyksilö tai joukon tulee olla homogeeninen, koska samalla yksilöllä ei voida suorittaa toistokokeita. Oikeastaanhan ei voida tietää, miten tietty yksilö reagoi lääkkeeseen. Vaikka kontrafaktuaalitapauksen todennäköisyys laskettaisiinkin, tulee välttää johtopäätösten vetäminen koskien tiettyä yksilöä.

Kontrafaktuaalitapahtuman todennäköisyys voidaan kuitenkin laskea rajoittumalla niiden ihmis- ja potilastyyppiparien avaruuteen, jotka pätevät käsillä olevassa tapauksessa [Pea00]. Jo tapahtunut tapahtuma määrää yksikäsitteisesti tämän avaruuden. Esimerkiksi jos tiedetään, että lääkettä suositeltiin, potilas otti lääkkeen ja parani, niin mahdolliset tyyppiparit ovat seuraavat: tottelevainen ja aina paraneva, aina lääkettä syövä ja lääkkeen avulla parantuva, aina lääkettä syövä ja aina paraneva sekä tottelevainen ja lääkkeen avulla paraneva. Nyt voidaan laskea olisiko henkilö parantunut, jos lääkettä ei olisi suositeltu. Pari tottelevainen ja lääkkeen avulla paraneva ei kelpaa, sillä silloinhan tiedämme jo varmasti ettei henkilö tule paranemaan kun lääkettä ei suositella. Eli formaalisti lasketaan todennäköisyys

$$P(Y_{z_0} = y_1 | y_1, x_1, z_1) = \frac{q_1 + q_2 + q_3}{P(y_1, x_1 | z_1)},$$

missä osoittajassa on yllä mainittujen hyväksyttyjen parien esiintymistodennäköisyydet. Merkinnällä Y_{z_0} tarkoitetaan satunnaismuuttujaa Y tapauksessa, jossa lääkettä ei suositella. Osoittajalle joudutaan nyt laskemaan rajat käyttämällä lineaarista optimointia samalla tapaa kuin tarkkojen rajojen yhteydessä, sillä eri ryhmien ko-

kojahan ei tunneta. Saatua todennäköisyyttä voidaan pitää mittarina paranemistodennäköisyydelle ilman lääkkeen suosittelua. Myös Gibbsin otantaa käyttämällä voidaan laskea kontrafaktuaalisten tapahtumien todennäköisyyksiä [ChP99]. Tällöin aiemmin esiteltyä mallia kuvassa 2 muutetaan siten, että solmun v_R ja $ACE(X \to Y)$ välinen suhde korvataan kyselyyn sopivalla deterministisellä funktiolla. Jos dataa on vähän, vaikuttaa priori tietenkin myös tässä tapauksessa voimakkaasti lopputulokseen. Priorin vaikutusta lopputulokseen voi arvioida käyttämällä useampaa eri prioria ja vertaamalla tuloksia keskenään [Pea00].

4.2 Mallin todenmukaisuuden testaaminen

Muuttujaa Z kutsutaan instrumenttimuuttujaksi. Jotta edellä johdetut tulokset lääkkeen vaikutukselle pätisivät, tulee muuttujan Z olla satunnaistettu ja riippumaton latenteista tekijöistä U. Tarkkojen rajojen epäyhtälöistä voidaan johtaa seuraavat ehdot [Pea00]:

$$\begin{aligned}
P(y_0, x_0|z_0) + P(y_1, x_0|z_1) &\leq 1, \\
P(y_0, x_1|z_0) + P(y_1, x_1|z_1) &\leq 1, \\
P(y_1, x_0|z_0) + P(y_0, x_0|z_1) &\leq 1, \\
P(y_1, x_1|z_0) + P(y_0, x_1|z_1) &\leq 1.
\end{aligned} \tag{5}$$

Ehdot seuraavat tarkat rajat määrittelevistä epäyhtälöistä 3 ja 4 sekä vaatimuksesta, että jokainen ylärajoista on suurempi kuin kaikki alarajat. Jos epäyhtälöryhmä 5 ei päde joidenkin aineistojen yhteydessä, syynä on joko lääkkeensuosittelumenetelmän suora yhteys havaittuun vaikutukseen tai lääkkeensuosittelumenetelmän ja latenttien tekijöiden keskinäinen riippuvuus. Edellisestä päästään eroon tehokkaalla lumelääkkeen käytöllä ja jälkimmäisestä parantamalla satunnaistamismenettelyä [Pea00].

Epäyhtälöryhmä 5 voidaan suoraviivaisesti laajentaa koskemaan useampiarvoisia muuttujia tai jatkuvia muuttujia. Tekemällä lisäoletuksia testihenkilöiden käyttäytymisestä saadaan tarkempia ehtoja instrumenttimuuttujien testaamiseen [Pea00].

5 Yhteenveto

Muunnetulla sokkotestillä voidaan päätellä lääkkeen vaikutus luotettavammin kuin perinteisen tapaan suoritetulla testillä. Muunnetun testin tapauksessa tiedetään varmemmin milloin testitulokset eivät ole tarkkoja. Epäyhtälöillä 2, 3 ja 4 määriteltyjen luonnollisten ja tarkkojen rajojen avulla saadaan lääkkeen vaikutukselle tietty väli, jolta todellinen arvo löytyy. Käyttämällä Gibbsin otantaa suurella aineistolla saadaan melko luotettava todennäköisyysjakauma keskimääräiselle kausaaliselle vaikutukselle. Mitä suurempi aineisto on sitä tiukemmin jakauma keskittyy oikean arvon ympärille tai oikean välin sisälle. Esitettyjen tulosten käyttäminen päättelyssä edellyttää, että koejärjestely noudattaa annettua kausaalimallia. Kausaalimallin ja havaintoaineiston yhteensopivuutta voidaan tarkastella yksinkertaisilla testeillä.

Lähteet

ChP99 Chickering, D. ja Pearl, J., A clinician's tool for analyzing noncompliance. Teoksessa *Computation, Causation, and Discovery*, Glymour, C. ja Cooper, G., toimittajat, AAAI Press/MIT Press, 1999, sivut 407–424.

Daw97 Dawid, A. P., Causal inference without counterfactuals. Tutkimusraportti, University College London, 1997. URL `http://www.ucl.ac.uk/Stats/research/psfiles/188.zip,9.4.2003`.

Pea95 Pearl, J., Causal inference from indirect experiments. *Artificial Intelligence in Medicine*, 7,6(1995), sivut 561–582. URL `citeseer.nj.nec.com/article/pearl95causal.html,9.4.2003`.

Pea00 Pearl, J., *Causality: models, reasoning, and inference.* Cambridge university press, Cambridge, 2000.