

Kandidaatin tutkielma
Kausaalipäättely ja valikoitumisharha

Riku Laine
Valtiotieteellinen tiedekunta
Helsingin yliopisto

21. huhtikuuta 2019

Sisältö

1	Kiitokset – Acknowledgements	2
2	Johdanto	3
2.1	”Kausaalipäätely uutena paradigmana”	3
2.2	Valikoitumisharha – seulotun aineiston ongelma	4
3	Aineiston generointi	6
4	Menetelmät	8
4.1	Aiemmat tutkimukset?	8
4.2	Validointimetodit	8
4.3	Verkkoteoria	8
4.4	Kausaalipäätely	10
4.4.1	Johdanto?	11
4.4.2	Merkinnät	11
4.4.3	Määritelmät	11
4.4.4	Malli	12
5	Tulokset	14
5.1	Synteettinen	14
5.2	Compas	14
6	Diskussio	15
	Lähteet	16

Luku 1

Kiitokset – Acknowledgements

Tämän tutkielman aikana on tullut esiin takuujärjestelmään liittyvät ongelmat ja sovel-
lusalueen yhteiskunnallinen merkitys. Tutkielman teko on ollut minulle erityisen miele-
kässtä antoisan aiheen ja mieleisten yhteistyökumppanien vuoksi. Olen kirjoittanut tämän
kandidaatintutkielman yhteistyössä Helsingin yliopiston tietojenkäsittelytieteen osaston
apulaisprofessorin Michael Mathioudakis ja tohtoritutkijan Antti Hyttisen kanssa. He
tarjosivat minulle aiheen ja merkittävää tukea sekä tärkeitä kommentteja tämän tutkiel-
man kirjoittamisen aikana.

Tämän tutkielman on tarkastanut XYZ. Haluan kiittää kaikkia edellä mainittuja hen-
kilöitä sekä ystäviäni ja perhettäni, jotka tukivat minua tämän tutkielman tekemisessä.

Helsingissä 21. huhtikuuta 2019,
Riku Laine

I would like to wholeheartedly thank assistant professor Michael Mathioudakis from Uni-
versity of Helsinki’s Department of Computer Science for numerous things. He provided me
this extremely interesting thesis topic and provided insightful and encouraging comments
throughout the process. Antti Hyttinen from the same department also gave important
insight in the causal modelling and commented on the content.

Luku 2

Johdanto

Tämän tutkielman tavoitteena on luoda kausaalipäätelyn avulla algoritmi, jolla voimme arvioida ennustavien mallien tarkkuutta, kun käytettävissä on ainoastaan valikoitumisharhasta kärsivää aineistoa. Samankaltaista asetelmaa ovat julkaisuissaan käsitelleet muun muassa Lakkaraju ja Madras [3, 4]. Pyrin tutkielmassani luomaan joustavamman ja tarkemman vaihtoehdon Lakkarajun luomalle supistusalgoritmille, mutta esitän ensin yleistä taustaa kausaalipäätelystä ja valikoitumisharhasta.

2.1 ”Kausaalipäätely uutena paradigmana”

Kuten Pearl ja Mackenzie esittävät kirjassaan Miksi, ihmisillä on luontainen kausaalisen päätelyn taito [7]. Tavalliset tilastollisen päätelyn menetelmät eivät tarjoa tapaa määritellä kausaalista yhteyttä: aineistosta voidaan päätellä erilaisia *korrelaatioita*, mutta päätely *A johtuu B:stä* vaatii uudenlaista lähestymistapaa. Käytännön tutkimuksessa kausaaliset yhteydet kiinnostavat erityisesti lääketieteen alalla. Kuten Kalisch toteaa, aiemmin kausaalisuuden päätely on perustunut korrelaatioiden havaitsemiseen. On hypotetisoitu, että biomarkkerin ja taudin samanaikainen ilmaantuminen viittaisi siihen, että markkeri aiheuttaa taudin. Voimmeko siis markkeria käsittelemällä vaikuttaa tautiin tai jopa parantaa se? [1]

Syy-seuraussuhteen matemaattinen määrittely vaatii uutta lähestymistä myös todennäköisyyslaskennan merkintöihin. Pearl käyttää alkuperäisessä, englanninkielisessä kirjallisuudessa merkintää 'do' ilmaisemaan interventiota. Merkinnällä halutaan erottaa tavanomainen ehdollinen todennäköisyys $\mathbb{P}(Y|X = x)$ interventiosta, jossa asetamme muutujan X arvoon x : $\mathbb{P}(Y|\text{do}(X = x))$. Kimmo Pietiläinen käyttää kirjan suomennoksessa do-operaattorista käännöstä *tee*, mutta seuraan tässä tutkielmassa Pearl'n merkintöjä, ellen erikseen muuta mainitse [7]. Alalla käytetään myös muita, alaindekseillä rikastettuja

merkintätapoja [6]. Esittelen käyttämäni merkinnät tarkemmin kappaleessa 4.4.2.

Kausaalipäätelyssä mallit voidaan esittää graafeina, eli verkkoina. Verkoista voidaan suoraan lukea eri muuttujien relaatiot kausaalisuuden suuntien ja riippuvuuksien suhteen.

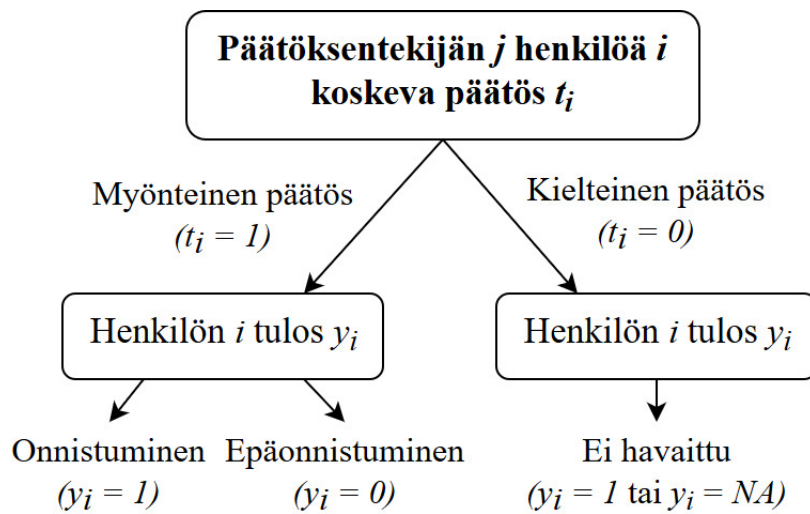
2.2 Valikoitumisharha – seulotun aineiston ongelma

Aineiston luova mekanismi on esitetty kuvassa 2.1 ja toimii siten, että aluksi jokin henkilö tai muu entiteetti saapuu päätöksentekijän eteen seulottavaksi. Päätöksentekijän tavoitteena on estää haitallinen tulos ($y = 0$) pitäen samalla myönteisten päätösten ($t = 1$) määrä mahdollisimman pienenä. Seuloja pyrkii siis antamaan kielteisen päätöksen kaikille niille, joilla epätoivottava tulos on todennäköisin. Päätöksen jälkeen henkilö siirtyy vaiheeseen, jossa Kohtalo määrittää hänelle tuloksen $y \in \{0,1\}$. Kielteisen päätöksen saaneille tulos voidaan merkitä puuttuvaksi tai onnistuneeksi, koska haitallista tapahtumaa ei havaita.

Aineiston generoivaa mekanismia voidaan havainnollistaa lääke- ja oikeustieteen alan esimerkillä. Henkilöllä viitataan ensin mainitussa potilaaseen ja jälkimmäisessä epäiltyyn. Seuloja voi olla esimerkiksi lääkäri, joka päättää annetaanko potilaalle vahvempaa ja samalla kalliimpaa lääkettä, jolloin relapsia ei havaita. Oikeudellisessa asetelmassa seulojalla voidaan tarkoittaa tuomaria, joka päättää epäillyn vapauttamisesta takuita vastaan ilman pelkoa rikoksen uusimisesta. Molemmilla päättäjillä on selkeä kannustin estää haitalliset tulokset – sairauskohtaukset tai rikokset – pitäen samalla päätöksistä aiheutuvat rasitteet yhteiskunnalle ja yksilöiden elämille mahdollisimman pienenä.

Havaintoja voi puuttua erilaisissa tutkimuksissa useista eri syistä. Kyselytutkimuksissa vastauskatoa voi syntyä esimerkiksi vastaajan haluttomuudesta vastata kysymykseen tai yksinkertaisesti siitä syystä, että vastaajaa ei tavoiteta. Jos aineiston puuttuneisuusmekanismi on luonteeltaan täysin satunnainen, eli vastauksen puuttuneisuus ei liity millään tavalla mitattuihin muuttujiin, voidaan sanoa aineistoa puuttuvan *täysin satunnaisesti*. Käänteisessä tapauksessa voidaan puhua *ei-satunnaisesta puuttuvuudesta*. [2]

Tässä tutkielmassa tarkasteltavassa asetelmassa havaintojen puuttuminen liittyy sekä havaittuihin että havaitsemattomiin muuttujiin. Puuttuneisuuden voidaan sanoa olevan *satunnaista ehdollisesti*, koska aineistoa puuttuu vain yksilöiltä, joilla on korkea todennäköisyys haitalliseen tulokseen. (Erilaisia aineiston puuttuneisuusmekanismeja esitellään laajemmin esimerkiksi Laaksosen kirjassa *Surveyymetodiikka*.) Puuttuneisuutta voidaan korvata imputoinnilla, jolla yritetään tehdä mahdollisimman hyvä arvaus puuttuvasta arvosta. Todistan tutkielmassani myöhemmin, että kausaalipäätelyä hyödyntämällä voimme estimoida havaitusta, valikoimusharhaisesta aineistosta haluttuja tunnusukuja ilman imputointia harhattomasti. [2] Englanninkielisessä kirjallisuudessa seulotun aineiston ongelmasta on alettu käyttää Lakkarajun esittämää termiä *selective labels* [3].



Kuva 2.1: Valikoitumisharha

Luku 3

Aineiston generointi

Synteettinen aineisto luotiin Lakkarajun artikkelissaan selostamalla tavalla [3]. Aineistoon simuloitiin kolme muuttujaa X , Z , ja W . Näistä muuttujista X vastaa informaatiota, joka on sekä mallin että tuomarin havaittavissa, eli informaatiota, joka on kirjattu oikeuden pöytäkirjoihin tai on kerättävissä muista rekistereistä, kuten vastaajan sukupuoli. Muuttujalla Z kuvataan tietoa, jonka vain tuomari voi havaita: kuten Lakkaraju havainnollistaa, tällaista voi olla esimerkiksi tieto siitä, onko vastaajalla perhettä mukana oikeussalissa [3]. W on mallissa havainnollistamassa reaali maailmaa. Muuttujalla esitämme aineistossa informaatiota, joka ei ole saatavilla päätöksentekijöille eikä mallille mutta vaikuttaa silti rikoksenuusimisriskiin. aineistossa nämä ovat kaikki riippumattomia standardinormaalijakautuneita satunnaismuuttujia, eli $X, W, Z \sim N(0, 1) \perp$.

Yhdistämme henkilöt satunnaisesti kuhunkin $M = 500$ tuomariin, joista jokaiselle määritellään hyväksymisprosentti $r \in [0, 1]$. Tuomarin hyväksymisprosentti määritetään ottamalla arvoja tasajakaumasta suljetulta väliltä $[0, 1; 0, 9]$ ja sitten pyöristämällä ne 10 desimaalin tarkkuuteen. Tulosuuttuja Y simuloidaan määrittämällä sen ehdollinen todennäköisyys seuraavasti: $\mathbb{P}(Y = 0 | X, Z, W) = \frac{1}{1 + \exp\{-(\beta_X X + \beta_Z Z + \beta_W W)\}}$, missä kertoimet β_X , β_Z ja β_W on asetettu arvoihin 1, 1 ja 0,2 vastaavassa järjestyksessä. [3]

Päätösmuuttujan T ehdollinen todennäköisyys $\mathbb{P}(T = 0 | X, Z) = \frac{1}{1 + \exp\{-(\beta_X X + \beta_Z Z)\}} + \epsilon$ missä $\epsilon \sim N(0, 0, 1)$ vastaa pientä määrää kohinaa. Henkilöltä i kielletään takuut, eli $T_i = 0$ jos muuttujan T ehdollinen todennäköisyys on tuomarin j suurimman $(1 - r) \cdot 100\%$ joukossa. Lopuksi koulutusaineisto suodatettiin siten, että saatavissa oli vain yksilöt, jotka päästettiin vapaaksi ($T = 1$). [3]

Muuttuja	Keskiarvo	Keskihajonta	Minimi	25%	50%	75%	Maksimi
acceptanceRate_R	0.48	0.23	0.10	0.26	0.47	0.65	0.89
X	-0.00	1.00	-4.66	-0.67	-0.00	0.67	3.83
Z	0.01	1.00	-4.85	-0.67	0.00	0.68	4.24
W	0.01	1.00	-4.03	-0.67	0.01	0.68	4.29
result_Y	0.50	0.50	0.00	0.00	0.00	1.00	1.00
probabilities_T	0.50	0.28	-0.34	0.28	0.50	0.72	1.30
decision_T	0.48	0.50	0.00	0.00	0.00	1.00	1.00

Taulukko 3.1: Synteettisen aineiston muuttujien hajontalukuja

Luku 4

Menetelmät

Tässä kappaleessa selostan analyyseissa, mallinnuksessa ja validoinnissa käyttämäni menetelmät.

4.1 Aiemmat tutkimukset?

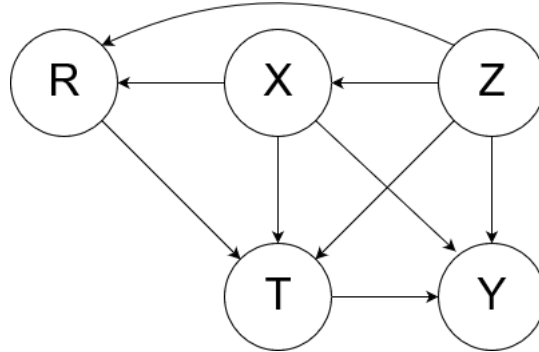
Aiemmat tutkimukset ovat lähestyneet monesta näkökulmasta, mutta ilman kausaatiota.

4.2 Validointimetodit

Tulosten arvioinnissa käytetään visuaalista tarkastelua ja XZY. Laskemme arvioista vapaaksi päässeiden uusijoiden suhteen kaikkiin tuomittuihin, eli niin sanotun virhesuhteen (failure rate).

4.3 Verkkoteoria

Kausaalipäättelyn mallit määritellään verkkoina. Esitän tässä kappaleessa lyhyesti kaikki tarvittavat verkkoteoreettiset määritelmät, joita tulen hyödyntämään. Noudatan määritelmissä Oinosta [5].



Kuva 4.1: Esimerkkiverkko $H = (V, E)$, missä $V = \{R, X, Z, T, Y\}$.

Määritelmä 4.1 (Suunnattu verkko). *Suunnattu verkko* G on pari (V, E) , missä $V \neq \emptyset$ on solmujen joukko ja

$$E = \{(a, b) \in V \times V \mid \text{solmusta } a \text{ on nuoli solmuun } b\}$$

on *kaarien* joukko.

Kuvassa 4.1 näkyvässä verkossa esimerkiksi $(X, R) \in E$, mutta $(T, Z) \notin E$, koska solmusta T ei ole nuolta solmuun Z . Lisäksi voidaan todeta, että kaarien joukkoon kuuluu yhdeksän järjestettyä paria ja solmujen joukko V käsittää viisi alkioa, jotka on lueteltu kuvatekstissä.

Määritelmä 4.2. Oletetaan, että $G = (V, E)$ on suunnattu verkko ja $a, b \in V$.

Merkintä $a \rightarrow b$ tarkoittaa, että $(a, b) \in E$. Tällöin sanotaan, että a on kaaren (a, b) *lähtösolmu* ja b on kaaren (a, b) *maalisolmu*. Sanotaan myös, että solmu b on solmun a *vierussolmu*.

Jos $(a, a) \in E$, sanotaan suunnatussa verkossa olevan *silmukka* solmussa a .

Esimerkkiverkossa H kaaren (Z, T) lähtösolmu on solmu Z ja maalisolmu solmu T . Lisäksi huomataan, että verkossa H ei ole yhtään silmukkaa.

Määritelmä 4.3 (Vierekkäisyys). Oletetaan, että $G = (V, E)$ on suunnattu verkko ja $a, b \in V$.

Jos solmujen a ja b välillä on nuoli, niin solmujen a ja b sanotaan olevan *vierekkäisiä*.

Kuvan 4.1 verkosta havaitaan, että melkein kaikki solmut ovat toistensa vierussolmuja. Ainoa poikkeus on solmut R ja Y , joiden välillä ei ole nuolta ja jotka eivät siten ole vierekkäisiä.

Määritelmä 4.4 (Yksinkertainen suunnattu verkko). Oletetaan, että $G = (V, E)$ on suunnattu verkko, jossa ei ole yhtään silmukkaa eli $(v, v) \notin E$ kaikilla $v \in V$.

Tällöin sanotaan, että G on yksinkertainen suunnattu verkko.

Esimerkkinä käytetystä verkosta H nähdään heti, että se on yksinkertainen suunnattu verkko, koska siinä ei ole yhtään silmukkaa. Yksinkertaisesta suunnatusta verkosta käytetään englanniksi nimitystä *directed acyclic graph* ja se saatetaan lyhentää DAG.

Määritelmä 4.5 (Polku ja suunnattu polku). Oletetaan, että G on yksinkertainen verkko ja $n \in \mathbb{N}, n \geq 1$.

Verkon G solmujen jono v_1, \dots, v_n on *polku* solmusta v_1 solmuun v_n , jos jonon jokaisesta solmusta on kaari jonon seuraavaan solmuun. Polkua voidaan merkitä $v_1 \rightsquigarrow v_n$.

Jos verkko G on suunnattu verkko, $a, b \in V$ ja kaikki polun $a \rightsquigarrow b$ kaaret kulkevat kaarien suuntien mukaisesti, voidaan täsmentää, että polku $a \rightsquigarrow b$ on *suunnattu polku*.

Huomataan, että esimerkkinä käytetyssä verkossa H on useita polkuja solmusta R solmuun Y . Polku $R \rightarrow T \rightarrow Y$ on suunnattu polku ja $R \leftarrow X \rightarrow Y$ on tavallinen polku, sillä solmujen R ja X välillä kuljetaan nuolen suunnan vastaisesti.

Määritelmä 4.6. Oletetaan, että $G = (V, E)$ on suunnattu verkko ja $a, b \in V$.

Jos on olemassa suunnattu polku $a \rightsquigarrow b$, niin solmun b sanotaan olevan solmun a *jälkeläinen*. Vastaavasti tällöin sanotaan solmun a olevan solmun b *vanhempi*.

Esimerkiksi kuvan 4.1 verkossa solmulla Y ei ole jälkeläisiä ja solmun Z jälkeläiset ovat kaikki muut verkon solmut pois lukien se itse, eli solmun Z jälkeläiset on joukko $V \setminus \{Z\}$.

4.4 Kausaalipäätely

Erityisesti [6]. Esittele merkunnät, määritelmät ja malli. Käännökset Miksi-kirjaa muokailleen?

4.4.1 Johdanto?

Kausaalipäättelyssä mallit määritellään usein yksinkertaisina suunnattuina verkkoina. Mallin määrittämistä verkosta voidaan suoraan lukea kausaaliset riippuvuussuhteet ja malliin kuuluvat muuttujat. Jos mallissa on solmut A ja B ja jos solmu B on solmun A jälkeläinen, niin muuttujalla A on mallin mukaan jonkinlainen kausaalinen vaikutus muuttujaan B . Jos verkossa muuttujien välillä ei ole jälkeläisyssuhdetta, niin ne ovat toisistaan riipumattomat. Kausalisen vaikutuksen funktionaalista muotoa ei usein määritellä.

* Usein funktionaalista muotoa ei määritellä, lisää tähän ne nuoliversiot yhtälöistä havainnollistamaan, että siirrytään yhtäsuuruudesta määräytymiseen [1]

4.4.2 Merkinnät

Kausaalipäättelyssä käytettävät merkinnät noudattelevat pitkälle tavallisia todennäköisyyslaskennan merkintöjä. Kun selvitetään muuttujan X vaikutusta muuttujaan Y ja tehdään interventio asettamalla muuttuja X arvoon x_0 , sitä merkitään $\mathbb{P}(Y|\text{do}(X = x_0))$.

4.4.3 Määritelmät

Määritelmä 4.7 (Takaovikriteeri, *back-door criterion*). Joukko \mathcal{S} sulkee / katkaisee (blocks) polun p , jos vähintään toinen seuraavista ehdoista on voimassa:

- (a) Polku p sisältää vähintään yhden solmun, joka on jonkin polun kulkusuuntaisen kaaren lähtösolmu ja kuuluu joukkoon \mathcal{S} . (arrow-emitting)
- (b) Polku p sisältää vähintään yhden käänteisen haarukkasolmun (collision node), joka ei kuulu joukkoon \mathcal{S} ja jolla ei ole jälkeläisiä joukossa \mathcal{S} .

Määritelmä 4.8. Oletetaan, että halutaan selvittää (satunnais)muuttujan X kausaalista vaikutusta muuttujaan Y . Joukko \mathcal{S} on *riittävä* tasoitukseen (adjustment), kun seuraavat ehdot ovat voimassa: **sufficient to adjustment = identifioituva?**

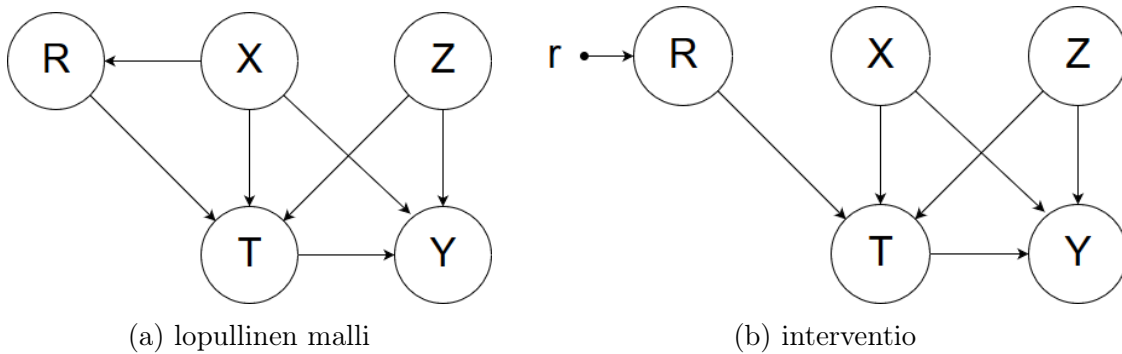
- (1) Yksikään joukon \mathcal{S} alkioista ei ole solmun X jälkeläinen.
- (2) Joukon \mathcal{S} alkiot katkaisevat kaikki määritelmän 4.7 mukaiset polut / ”takaovireitit” solmusta X solmuun Y .

4.4.4 Malli

Mallimme määrittelevä graafi on kuvattuna alla kuvassa 4.2a. Mallissamme on viisi muuttujaa, jotka on esitelty taulukossa 4.1. Muuttujalla R kuvataan vapautusprosenttia, eli sitä prosentuaalista osuutta henkilöistä, joilla on suurin vaara ja joilla silti annetaan positiivinen päätös. X ilmentää henkilön henkilökohtaisia ominaisuuksi, COMPAS-aineiston tapauksessa esimerkiksi henkilön ikää ja sukupuolta. Lisäksi muuttujan X tiedot on sekä tuomarin että ennustavan mallin havaittavissa. Muuttuja Z on muuttuja, jonka tuomari tai muu asiantuntija voi havaita, mutta joka on mallin ulottumattomissa. Muuttujan Z voidaan ajatella esimerkiksi oikeuskäsittelyjen tapauksessa kuvaavan onko henkilöllä perhe mukana oikeussalissa.

Muuttuja	Kuvaus
R	Vapautusprosentti, vapautumiskynnys
X	Henkilökohtaiset muuttujat, kirjalliset
Z	Henkilökohtaiset muuttujat, päättäjän havaitsemat
Y	Uusinta, $Y = 0$ uusi, 1 niin ei uusinut
T	0 on jail, 1 on bail

Taulukko 4.1: Mallin muuttujien selitteet



Kuva 4.2: Kausaalimallit graafina

Algoritmi 1 Kausaalialgoritmi

Syöte: aineisto $(\mathbf{x}, t, y) \in \mathcal{D}_t, \mathcal{D}_v$ ja hyväksymisaste $r \in [0, 1]$, missä \mathcal{D}_t on testiaineisto ja \mathcal{D}_v validointiaineisto.

Tuloste: $\mathbb{P}(Y = 0 | \text{do}(R = r))$

- 1: Määritä $f(x) = \mathbb{P}(X = x)$ testiaineistosta.
 - 2: Ennusta vastetta Y selittävillä muuttujilla X käyttäen harjoitusaineiston havaintoja, joilla $T = 1$.
 - 3: Määritä harjoitusaineiston jokaiselle havainnolle $P(Y = 0 | X = x)$ käyttäen yllä olevaa mallia.
 - 4: Järjestä havainnot nousevaan järjestykseen edellisen kohdan todennäköisyyksien mukaan.
 - 5: Alusta muuttuja **summa** = 0.
 - 6: **for all** Jokaiselle parametriavaruuden pisteelle **do**
 - 7: $p_x \leftarrow P(X = x)$
 - 8: $\mathcal{D}_{\S} \leftarrow \{\mathcal{D} | X = x\}$
 - 9: Assign first $r \cdot 100\%$ observations from \mathcal{D}_{\S} to \mathcal{D}_{rx}
 - 10: $p_t \leftarrow \frac{|\{\mathcal{D}_{rx} | T = 1\}|}{|\mathcal{D}_{rx}|}$
 - 11: $\mathcal{D}_{tx} \leftarrow \{\mathcal{D}_x | T = 1\}$
 - 12: $p_y \leftarrow \frac{|\{\mathcal{D}_{tx} | Y = 0\}|}{|\mathcal{D}_{tx}|}$
 - 13: Lisää muuttujaan **summa** tulo $p_y \cdot p_t \cdot p_x$
 - 14: **end for**
 - 15: **return** **summa**
-

Luku 5

Tulokset

5.1 Synteettinen

5.2 Compas

Luku 6

Diskussio

```
# R-koodi, tulos sama
library(igraph)
library(causaleffect)
# simplify = FALSE to allow multiple edges
g <- graph.formula(X -+ R, X -+ D, X -+ Y, R -+ D , D -+ Y, D -+ Y, Y -+ D, simplify = FALSE)
# Here the bidirected edge between X and Z is set to be unobserved in graph g
# This is denoted by giving them a description attribute with the value "U"# The edges
g <- set.edge.attribute(graph = g, name = "description", index = c(6,7), value = "U")

res <- causal.effect("Y", "R", G = g)
```


Lähteet

- [1] Kalisch, Markus ja Peter Bühlmann: *Causal structure learning and inference: a selective review*. Quality Technology & Quantitative Management, 11(1):3–21, 2014.
- [2] Laaksonen, Seppo: *Surveyymetodiikka: Aineiston kokoamisesta puhdistamisen kautta analyysiin*. bookboon.com, 2 painos, 2013.
- [3] Lakkaraju, Himabindu, Jon Kleinberg, Jure Leskovec, Jens Ludwig ja Sendhil Mullainathan: *The Selective Labels Problem: Evaluating Algorithmic Predictions in the Presence of Unobservables*. Teoksessa *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '17, sivut 275–284, New York, NY, USA, 2017. ACM, ISBN 978-1-4503-4887-4. <http://doi.acm.org.libproxy.helsinki.fi/10.1145/3097983.3098066>.
- [4] Madras, David, Elliot Creager, Toniann Pitassi ja Richard Zemel: *Fairness Through Causal Awareness: Learning Latent-Variable Models for Biased Data*. arXiv preprint arXiv:1809.02519, 2018.
- [5] Oinonen, Lotta: *Johdatus yliopistomatematiikkaan*, Tammikuu 2016. Samannimisen kurssin kurssimateriaali.
- [6] Pearl, Judea: *An introduction to causal inference*. Int J Biostat, 6(2):Artikkeli 7, Helmikuu 2010. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2836213/>.
- [7] Pearl, Judea ja Dana Mackenzie: *Miksi : syyn ja seurauksen uusi tiede*. Terra Cognita, Helsinki, 2018, ISBN 978-952-5697-93-3. Suomennos Kimmo Pietiläinen.