UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO
Matematika - 2. stopnja

Veronika Pegan

Pogoji za ugotavljanje interakcij v zadostnih kavzalnih modelih z binarnimi spremenljivkami

Magistrsko delo

Mentor: doc. dr. Jaka Smrekar

Ljubljana, 2013

Podpisana Veronika Pegan izjavljam:

- da sem magistrsko delo z naslovom *Pogoji za ugotavljanje interakcij v zadostnih kavzalnih modelih z binarnimi spremenljivkami* izdelala samostojno pod mentorstvom doc. dr. Jaka Smrekarja in
- da Fakulteti za matematiko in fiziko Univerze v Ljubljani dovoljujem objavo elektronske oblike svojega dela na spletnih straneh.

Ljubljana, 1. 9. 2013 Podpis: ....

Kazalo

Program dela

V magistrskem delu študirajte pogoje za ugotavljanje interakcij v zadostnih kavzalnih modelih z binarnimi (dihotomnimi) spremenljivkami.

Temeljna literatura:

- T.J. Vanderweele, T.S. Richardson, *General theory for interactions in sufficient cause models with dichotomous exposures*, The Annals od Statistics **40**(4) (2012) 2128-2161.
- T.J. Vanderweele, J.M. Robins, *Empirical and counterfactual conditions for sufficient cause interactions*, Biometrika **95**(1) (2008) 49-61.

Ljubljana, maj 2013

Mentor:
doc. dr. Jaka Smrekar

Povzetek

V nalogi obravnavamo kavzalni model zadostnih komponent s pomočjo teorije možnih izidov, v kateri so dogodki binarne slučajne spremenljivke. Definiramo zadostne vzroke, minimalne zadostne vzroke ter odločilne in neodvečne množice zadostnih vzrokov. S pojmom odločilnih množic zadostnih vzrokov zaobjamemo idejo zadostnega in potrebnega vzroka. Definiramo ireducibilne komponente kot potrebne dele zadostnega vzroka ter jih karakteriziramo in primerjamo z minimalnimi zadostnimi vzroki. Oblikujemo empirične testi ireducibilnosti in minimalnosti zadostnih vzrokov najprej brez pogoja monotonosti, potem pa še pri pogoju monotonosti. Formulacija testov zahteva odsotnost zavajajočega učinka, ki je prisoten v obliki interakcij med določenimi dogodki. Zavajajoč učinek podrobneje spoznamo pri računanju kavzalnega učinka. Definiramo singularne minimalne zadostne vzroke kot enolične minimalne zadostne vzroke in obravnavamo empirične teste singularnosti. Posebej spoznamo teste ireducibilnosti, minimalnosti in singularnosti v nasičenih linearnih Bernoullijevih regresijskih modelih. Nazadnje se osredotočimo na kavzalno zvezo med genotipi, kajenjem in rakom sečnega mehurja.

Abstract

In this paper we consider the sufficient cause model using the potential outcome framework. Events are binary random variables. We define sufficient causes, minimal sufficient causes, determinative and nonredundant sets of sufficient causes. With the notion of determinative set of sufficient causes we embrace the idea of a sufficient and necessary cause. We define irreducible components as necessary components of a sufficient cause, we characterize them and compare them with minimal sufficient causes. We design empirical tests for irreducibility and for a minimal sufficient cause first without monotonicity and then under monotonicity. The tests require the absence of the confounding effect, which may be present in the form of an interaction between specific events. We get to know the confounding effect further in the calculation of the causal effect. We define singular minimal sufficient causes as unique minimal sufficient causes and deal with empirical tests for singularity. In particular we consider tests for irreducibility, minimality and singularity in saturated linear Bernoulli regression models. Finally, we focus on the causal relationship between the genotypes, smoking and bladder cancer.

Math. Subj. Class. (2010): 62D05, 62F99, 62J12, 92D30
Ključne besede: zadostni kavzalni model, model možnih izidov, potreben in zadosten vzrok, interakcije, Bernoullijev regresijski model
Keywords: sufficient cause model, potential outcome model, necessary and sufficient cause, interactions, Bernoulli regression model

1 Uvod

Problem, ki stoji v ozadju sledeče magistrske naloge, je vprašanje kavzalnosti in njene obravnave. Kavzalnost je zveza med dvema dogodkoma, kjer je drugi dogodek posledica prvega. Prvi dogodek imenujemo *vzrok*, drugega pa *učinek*.

Narava vzrokov in učinkov otežuje raziskovanje kavzalnosti predvsem iz dveh razlogov. Prvi sestoji v tem, da je vzrok v večini primerov sestavljen iz večjega števila dejavnikov. Ti so lahko neposredni ali pa učinkujejo preko drugih posrednih dejavnikov. Drugi razlog je, da vzroki in posledice niso le spremembe ali dogodki, kot jih običajno opredeljujemo, ampak tudi predmeti, procesi, lastnosti, spremenljivke in dejstva.

Kavzalne zveze sklepamo iz opazovanih podatkov in eksperimentov. Poslužujemo se številnih matematičnih modelov in teorij. Modele, ki zaobjamejo kavzalno zvezo med dvema dogodkoma, imenujemo *kavzalni modeli*.

Metode zbiranja podatkov imajo tudi pomembno vlogo pri kavzalnemu sklepanju. Idealna metoda zbiranja podatkov je naključno določanje vzorca. Nenaključni podatki lahko povsem popačijo zaključke. Pri tem moramo tudi paziti na omejitve, ki jih vključujejo eksperimenti. Namreč, nekatere eksperimente ne moremo izvesti zaradi moralnih, pravnih ali drugih razlogov. Nekateri eksperimenti pa so povsem fizično nemogoči. To so na primer raziskave glede rase in spola, saj jih omejujejo predsodki in podzavestne mentalne pregrade. Na primer, učinek rase ali spola na delovno produktivnost ne moremo izračunati.

Filozofija se z vprašanjem kavzalnosti ubada že tisočletja. V zahodni filozofiji razprava sega vsaj do Aristotela in ostaja aktualna tematika v sodobni filozofiji.

Vrste vzrokov. V splošnem vzroke razlikujemo v dve različni vrsti: potrebne in zadostne vzroke. Če je dogodek X *potreben vzrok* učinka Y, potem prisotnost Y-a nujno pomeni prisotnost X-a. Prisotnost X-a pa ne pomeni, da se bo Y pripetil.

Če pa je X *zadosten vzrok*, potem prisotnost X-a pomeni prisotnost Y-a. Prisotnost Y-a pa ne pomeni prisotnost X-a, saj ima morda Y nek drugi vzrok Z.

Pojem vzroka iz vsakdanjega življenja se običajno nanaša na potrebno komponento zadostnega vzroka. Poglejmo primer zažiga hiše. V vsakdanjem življenju lahko pojmujemo kratek stik kot vzrok zažiga hiše. Kratek stik pa ni zadosten vzrok sam po sebi, ampak je del zadostnega vzroka, ki je sestavljen iz množice naslednjih dogodkov: kratek stik, bližina vnetljivega materiala in odsotnost gasilcev. Ta množica dogodkov ni potreben vzrok, saj obstajajo številne druge množice, ki povzročijo zažig hiše. Kratek stik je potrebna komponenta zgornje množice, saj v njegovi odsotnosti ne bi prišlo do zažiga hiše, če bi bilo v bližini veliko vnetljivega materiala in ne bi bilo gasilcev. Torej niso vzroki iz vsakdanjega življenja zadostni vzroki, ampak so le potrebne komponente zadostnih vzrokov.

Teorije. Kavzalnost lahko obravnavamo s pomočjo različnih teorij. V sledeči magistrski nalogi bomo uporabljali *teorijo možnih izidov*, ki poudarja pomembnost tistih podatkov, ki jih ne moremo opazovati in jih ne bomo nikoli mogli pridobiti. To so podatki o učinku ob odsotnosti vzroka. Točneje, teorija upošteva podatke o učinku, ki bi se pripetil, ko bi se vzrok ne pripetil. Vedno lahko opazujemo le učinek, ki se pripeti po določenem vzroku in ne moremo nikoli natančno vedeti, kaj bi se pri istih okoliščinah pripetilo v odsotnosti vzroka. Posledično ne moremo natančno opredeliti kavzalne zveze. Pomanjkanju takih podatkov pravimo *osnovni problem kavzalnega sklepanja*. Namreč vrnitev v trenutek, ko se je pripetil vzrok, in izvedba eksperimenta v odsotnosti vzroka sta neizvedljiva. Teorija možnih izidov obravnava učinke ob odsotnosti vzroka kot neznane spremenljivke.

Nekatere teorije opisujejo kavzalne zveze s pomočjo grafov. Vozlišča predstavljajo dogodke, in usmerjene povezave pa kavzalne zveze. Pomemben doprinos pri preučevanju kavzalnosti s pomočjo grafov je prispeval ameriški računalničar in filozof Pearl Judea izraelskega porekla.

Obstaja še veliko drugih teorij, ki obravnavajo kavzalnost kot na primer analiza poti in modeli strukturnih enačb, ki ne bodo predmet naše obravnave.

Področja. Raziskovanje kavzalnih zvez je predmet velikega zanimanja na številnih področjih.

V socioloških in družboslovnih študijah so predmet zanimanja naslednje kavzalne zveze: vpliv velikosti razreda na ocene dijakov; posledice uvedbe denarnih nagrad pri šolskem uspehu za študente, ki bi najverjetneje šolo zapustili; vpliv zgodnjega začetka šolanja na šolske dosežke; vpliv bolnic na zdravje; vpliv izobraževalnih programov za dolgotrajno brezposelne ljudi na zaposljivost in plačo; učinek prostovoljnega vojaškega dela na bodočo plačo; vpliv družine na izobraževanje; vpliv katoliških šol na učenje itd.

Na ekonomskem področju raziskujejo vpliv povpraševanja na ponudbo, vpliv šolanja na plačo, učinke kolonializma na ekonomsko rast, vpliv vladne strukture na ekonomsko rast itd.

V epidemiologiji obravnavajo kavzalno zvezo med zdravljenjem in sledečim zdravstvenim stanjem. Raziskovanje na tem področju je ključnega pomena, saj odprta vprašanja o kavzalnih zvezah povzročajo napačne interpretacije in zahteve po slabih ukrepih na področju javnega zdravstva. Iz njih pa sledijo negativne pravne in finančne posledice. Kočljiva kavzalna zveza je na primer vpliv splava na rakava obolenja na dojki. Vprašanje je bilo predmet obravnave v številnih raziskavah in so šele pred kratkim nekatere študije neizpodbitno pokazale, da je vzročna povezava resnična [5]. Tudi vzročna povezava med kajenjem in rakom na pljučih je bila vprašljiva v letih 1940 in 1950. Odprto vprašanje je bilo vir številnih neutemeljenih zaključkov. Potrdili so jo šele leta 1954. Primer kočljive kavzalne zveze je tudi negativen vpliv hormonskih terapij na simptome menopavze. Raziskave so v nasprotju s pričakovanji pokazale neučinkovitost in negativne posledice zdravljenja s pomočjo hormonskih terapij.

V genetiki je raziskovanje usmerjeno v vpliv genetskih faktorjev na zdravstveno stanje osebkov. Raziskuje se kavzalno zvezo med genetskimi lastnostmi in raznimi obolenji. Nekatere genetske lastnosti namreč imajo posledice na dovzetnost imunskega sistema. V zadnjem poglavju bomo obravnavali primer iz genetike, točneje kavzalno zvezo med genotipi, kajenjem in rakom sečnega mehurja.

V sledeči magistrski nalogi bomo najprej predstavili model možnih izidov. Potem bomo obravnavali kavzalni model zadostnih komponent, kjer bomo spoznali pojme minimalnega zadostnega vzroka ter odločilne in neodvečne množice zadostnih vzrokov. V četrtem poglavju bomo definirali ter karakterizirali ireducibilne komponente in jih primerjali z minimalnimi zadostnimi vzroki. V petem poglavju bomo obravnavali teste ireducibilnosti in minimalnosti zadostnih vzrokov najprej brez pogoja monotonosti, potem pa še pri pogoju monotonosti. Formulacija testov zahteva odsotnost zavajajočega učinka, ki je prisoten v obliki interakcij med določenimi dogodki. V šestem poglavju bomo obravnavali še singularne minimalne zadostne vzroke in teste singularnosti. V sedmem poglavju bomo spoznali pogoje ireducibilnosti, minimalnosti in singularnosti v nasičenem linearnem Bernoullijevem regresijskem modelu. V zadnjem poglavju se bomo osredotočili na kavzalno zvezo med genotipi, kajenjem in rakom sečnega mehurja: uporabili bomo teorijo iz prejšnjih poglavij in pokazali, da niso domnevne interakcije statistično pomembne.

2 Model možnih izidov

V sledečem poglavju bomo predstavili *model možnih izidov* (potential outcome model), ki je poznan tudi pod imenom *Neyman-Rubinov model kavzalnosti* ali *model hipotetičnih izidov* (counterfactual model).

2.1 Osnovna ideja

Poglejmo najpreprostejšo formulacijo modela možnih izidov, kjer obravnavamo en sam vzrok. Populacijo osebkov razdelimo v dve skupini: *zdravljeno skupino* in *kontrolno skupino*. Osebki v zdravljeni skupini so na zdravljenju, osebki v kontrolni skupini pa ne uživajo nobenih zdravil, oziroma uživajo placebo. Populacijo označimo z Ω, posameznega osebka pa z $\omega \in \Omega$.

Naj bo X slučajna spremenljivka definirana za vse osebke v populaciji, ki označuje izpostavljenost zdravljenju oziroma kontroli in zavzema vrednosti iz $\{0, 1\}$. Točneje, $X(\omega) = 1$, če je osebek ω podvržen zdravljenju, ter $X(\omega) = 0$, če je osebek ω v kontrolni skupini. Izpostavljenost zdravljenju je običajno odvisna od odločitve posameznega osebka; lahko pa je naključno dodeljena ali pa jo določa zunanji opazovalec.

Naj bosta D_1 ter D_0 slučajni spremenljivki definirani za vse osebke v populaciji, ki označujeta zdravstveno stanje po zdravljenju oziroma kontroli. Točneje, $D_1(\omega)$ označuje zdravstveno stanje osebka ω po zdravljenju in $D_0(\omega)$ označuje zdravstveno stanje osebka ω po kontroli. Imenujemo ju *možna* ali *hipotetična izida* osebka ω. Predpostavimo, da obstajata oba možna izida nekega osebka, $D_1(\omega)$ in $D_0(\omega)$, zato sta spremenljivki D_0 in D_1 definirani na celotni populaciji. Model temelji na dejstvu, da dopuščamo obstoj vseh možnih izidov, ki jih dejansko ne moremo opazovati, ampak jih obravnavamo kot neznane izide definiranih spremenljivk.

Naj bo D slučajna spremenljivka definirana za vse osebke v populaciji, ki označuje zdravstveno stanje. Torej $D(\omega)$ označuje zdravstveno stanje osebka ω in velja:

$$D = XD_1 + (1 - X)D_0. \tag{2.1}$$

Kavzalni učinek zdravljenja na individualni ravni običajno definiramo

$$\delta(\omega) = D_1(\omega) - D_0(\omega).$$

Lahko ga pa tudi definiramo kot $D_1(\omega)/D_0(\omega)$ ali drugače. *Osnovni problem kavzalnega sklepanja* je dejstvo, da pri kontrolni skupini ne moremo opazovati možnega izida po zdravljenju in pri zdravljeni skupini ne moremo opazovati možnega izida v kontrolnem stanju. Posledično ne moremo izračunati kavzalnega učinka na individualni ravni. Kljub temu, lahko izračunamo številne povprečne kavzalne učinke populacije ali posameznih podskupin. *Povprečni učinek zdravljenja* označimo

$$\begin{aligned} E[\delta] &= E[D_1 - D_0] \\ &= E[D_1] - E[D_0]. \end{aligned}$$

Pogosto nas zanima tudi dvoje pogojnih povprečnih učinkov zdravljenja: *povprečni učinek zdravljenja na zdravljene osebke* oziroma na tiste, ki se običajno odločijo za zdravljenje (Average Treatment effect for the Treated)

$$\begin{aligned} E[\delta \mid X = 1] &= E[D_1 - D_0 \mid X = 1] \\ &= E[D_1 \mid X = 1] - E[D_0 \mid X = 1] \end{aligned}$$

ter *povprečni učinek zdravljenja za kontrolne osebke* oziroma tiste osebke, ki se običajno ne odločijo za zdravljenje (Average Treatment effect for the Controled)

$$\begin{aligned} E[\delta \mid X = 0] &= E[D_1 - D_0 \mid X = 0] \\ &= E[D_1 \mid X = 0] - E[D_0 \mid X = 0]. \end{aligned}$$

2.2 Posplošitev modela možnih izidov

Naj bo Ω populacija osebkov in $\omega \in \Omega$. Na vsak osebek vpliva s različnih dogodkov. Naj bodo $X_1, \ldots, X_s$, slučajne spremenljivke definirane za vse osebke v populaciji, ki te dogodke predstavljajo in zavzemajo vrednosti iz $\{0, 1\}$. Imenujemo jih *vzroki*. Označimo $X_i(\omega) = 1$, če se dogodek X_i vezan na osebka ω pripeti, ter $X_i(\omega) = 0$, če se dogodek X_i vezan na osebka ω ne pripeti.

Če velja $X_i = x_i$, pravimo, da dogodek X_i *zavzame vrednost* x_i iz $\{0, 1\}$ ter če velja $X_i = x_i$ za vsak $i \in \{0, \ldots, s\}$, pravimo, da vektor dogodkov $\mathbf{X} = (X_1, \ldots, X_s)$ zavzame vrednost $\mathbf{x} = (x_1, ..., x_s)$ iz $\{0, 1\}^s$.

Pri vsaki vrednosti vektorja $(x_1, ..., x_s)$ iz $\{0,1\}^s$ označimo z $D_{x_1,...,x_s}$ slučajno spremenljivko definirano za vse osebke v populaciji, ki označuje zdravstveno stanje, potem ko je vektor dogodkov $\mathbf{X} = (X_1, \ldots, X_s)$ zavzel vrednost $\mathbf{x} = (x_1, ..., x_s)$. Točneje, označimo z $D_{x_1,...,x_s}(\omega)$ zdravstveno stanje osebka ω, potem ko je vektor dogodkov $\mathbf{X} = (X_1, \ldots, X_s)$ zavzel vrednost $\mathbf{x} = (x_1, ..., x_s)$. Vrednost $D_{x_1,...,x_s}(\omega)$ pri določenem vektorju $(x_1, ..., x_s)$ imenujemo *možni* ali *hipotetični izid*.

Označimo z D slučajno spremenljivko definirano za vse osebke v populaciji, ki označuje zdravstveno stanje, torej je $D(\omega)$ zdravstveno stanje osebka ω. Vrednost $D(\omega)$ imenujemo *izid* pri osebku ω.

Naj bo $\mathbf{C} = \{X_1, \ldots, X_s\}$ in $\mathbf{c} = \{x_1, \ldots, x_s\}$. Izmenično bomo uporabljali oznake $D_{x_1,...,x_s}(\omega)$, $D_{X_1=x_1,...,X_s=x_s}(\omega)$, $D_{\mathbf{c}}(\omega)$ in $D_{\mathbf{C}=\mathbf{c}}(\omega)$. Vsak osebek ω v populaciji je vezan na 2^s možnih izidov: vsaka vrednost slučajnega vektorja $\mathbf{C} = (X_1, \ldots, X_s)$ porodi en možni izid. *Množico možnih izidov osebka* ω označimo z $\mathcal{D}(\mathbf{C}; \omega)$:

$$\mathcal{D}(\mathbf{C}; \omega) = \{D_{X_1=x_1,...,X_s=x_s}(\omega); (x_1, \ldots, x_s) \in \{0,1\}^s\}.$$

Množico možnih izidov populacije Ω pa označimo z $\mathcal{D}(\mathbf{C}; \Omega)$:

$$\mathcal{D}(\mathbf{C}; \Omega) = \{D_{X_1=x_1,...,X_s=x_s}(\omega); (x_1, \ldots, x_s) \in \{0,1\}^s, \omega \in \Omega\}.$$

Izide in možne izide pri osebku ω povezuje naslednji *aksiom doslednosti*:

$$D_{X_1=X_1(\omega),...,X_s=X_s(\omega)}(\omega) = D(\omega). \tag{2.2}$$

V nadaljevanju bomo obravnavali le kavzalne modele, kjer je spremenljivka izidov D binarna slučajna spremenljivka, torej zavzema le vrednosti iz $\{0,1\}$. Predstavljajmo si, da izid $D(\omega) = 1$ označuje prisotnost bolezni pri osebku ω, $D(\omega) = 0$ pa označuje odsotnost bolezni pri osebku ω.

2.3 Predpostavke modela možnih izidov

Model možnih izidov temelji na dveh predpostavkah, ki sledijo neposredno iz formulacije opisane v prejšnjem razdelku. Prva predpostavka zahteva, da so možni izidi pri vsakem osebku deterministični. Namreč, če množica dogodkov $\mathbf{C}$ zavzame vrednost $\mathbf{c}$ pri določenem osebku ω, potem ima osebek ω točno določeno zdravstveno stanje $D_{\mathbf{c}}(\omega)$.

Drugo predpostavko imenujemo *predpostavka o stabilni vrednosti zdravljenja ene enote* (Stable Unit Treatment Value Assumption) ali *predpostavka o odsotnosti motenj* (no interference assumption). Ta zagotavlja dobro definiranost možnih izidov in zahteva, da niso možni izidi določenega osebka pod vplivom drugih osebkov. Točneje možni izidi določenega osebka ne smejo biti niti pod vplivom vzrokov drugih osebkov niti pod vplivom možnih izidov drugih osebkov.

Primer neveljavnosti predpostavke SUTVA je poskus cepljenja, kjer pride do t. i. pojava *imunitete črede* (herd immunity), kjer cepljenje pomembnega dela populacije zaščiti tudi del populaciji, ki ni direktno razvil imuniteto. V tem primeru ni mogoče definirati funkcije D kot v prejšnjem razdelku.

Primer 2.1 *Poglejmo primer v modelu možnih izidov. Recimo, da je populacija sestavljena iz dveh osebkov. Naj bodo D, X_1, X_2 in X_3 binarne slučajne spremenljivke definirane na populaciji dveh osebkov, ki zavzemajo vrednosti iz $\{0,1\}$. Spremenljivke X_1, X_2 in X_3 predstavljajo vzroke, spremenljivka D pa izide. V levem delu tabele 1 si lahko ogledamo primer možnih izidov, v desnem pa primer dejanskih izidov.*

Tabela 1: *Možni in dejanski izidi v modelu s tremi vzroki in populacijo dveh osebkov.*

Osebek	$\mathbf{D_{000}}$	$\mathbf{D_{001}}$	$\mathbf{D_{010}}$	$\mathbf{D_{011}}$	$\mathbf{D_{100}}$	$\mathbf{D_{101}}$	$\mathbf{D_{110}}$	$\mathbf{D_{111}}$	$(\mathbf{X_1}, \mathbf{X_2}, \mathbf{X_3})$	$\mathbf{D}$
1	0	1	1	0	0	1	1	0	$(1,0,1)$	1
2	0	1	1	0	0	1	1	1	$(0,0,0)$	0

3 Zadostni kavzalni model

V sledečem poglavju bomo definirali zadosten vzrok, minimalen zadosten vzrok ter odločilno in neodvečno množico zadostnih vzrokov. Spoznali bomo, da pravimo odločilni množici zadostnih vzrokov *zadostni kavzalni model.* Zadosten vzrok je množica dogodkov, ki povzročijo neko posledico.

Model, ki obravnava *vzrok* kot množico dejavnikov, je predlagal epidemiolog Ken Rothman že leta 1976. Poudaril je, da je *vzrok* določene bolezni sestavljen iz številnih dejavnikov, iz katerih lahko sledi bolezensko stanje. Obenem je poudaril, da se dejavniki, ki sestavljajo vzrok, razlikujejo od osebka do osebka. Opredelil je *zadosten vzrok* kot množica dejavnikov, iz katerih nujno sledi bolezensko stanje.

Primer 3.1 *Poglejmo primer zadostnega vzroka, ki povzroči AIDS. Obravnavala sta ga epidemiologa Aschengrau in Seage. Komponente nekega zadostnega vzroka za AIDS so naslednje: kontakt s posameznikom okuženim z virusom HIV, tvegano spolno vedenje s posameznikom okuženim z virusom HIV in odsotnost zdravil proti retrovirusom, ki zmanjšujejo virusno breme HIVa.*

Prisotnost teh komponent pri določenem osebku bo nujno privedel do obolenja z AIDSom. Če je katera izmed komponent odsotna, ni nujno, da bo prišlo do bolezni.

Primer 3.2 *Poglejmo tri primere zadostnih vzrokov, ki povzročijo tuberkulozo. Zadosten vzrok tuberkuloze nekega osebka je množica naslednjih dejavnikov: slaba prehrana, gneča, slaba ventilacija, izpostavljenost tuberkulozi in odsotnost cepiva proti tuberkulozi. Zadosten vzrok nekega drugega osebka je sestavljen iz podmnožice prejšnjega zadostnega vzroka oziroma iz naslednjih dogodkov: gneča, slaba ventilacija, izpostavljenost tuberkulozi in odsotnost cepiva proti tuberkulozi. Zadosten vzrok tretjega osebka pa je na primer sestavljen iz slabe prehrane, izpostavljenosti tuberkulozi in AIDS-a. V vseh navedenih primerih zadostnega vzroka tuberkuloze je prisotna izpostavljenost tuberkulozi. Sama izpostavljenost tuberkulozi ni zadosten vzrok bolezni, vendar je pomembna komponenta, brez katere ne bi prišlo do bolezni, če bi ostale komponente bile še vedno prisotne. Komponenta je torej potrebna, vendar ni zadostna.*

3.1 Oznake in osnovne definicije

Bernoullijeve slučajne spremenljivke površinsko imenujemo *dogodki.* Označimo jih z velikimi tiskanimi črkami. Množice dogodkov označimo z velikimi krepkimi tiskanimi črkami: $\mathbf{C} = \{X_1, \ldots, X_s\}$. Vrednost dogodka iz $\{0, 1\}$ označimo z mali tiskanimi črkami x. Z zapisom $\mathbf{C} = \mathbf{c}$, kjer je $\mathbf{c} = (x_1, \ldots, x_s)$ vektor vrednosti iz $\{0, 1\}^s$, površno označimo, da je vektor $(X_1, \ldots, X_s)$ enak vektorju $(x_1, \ldots, x_s)$. V tem primeru pravimo, da množica $\mathbf{C}$ *zavzame vrednost* $\mathbf{c}$. Vektor enic označimo z $\mathbf{1}$, vektor ničel pa z $\mathbf{0}$. Kardinalnost množice $\mathbf{C}$ označimo z $|\mathbf{C}|$. Množice množic dogodkov pa označimo z gotsko pisavo: $\mathfrak{B} = \{\mathbf{C}_1, \ldots, \mathbf{C}_n\}$.

Dejanski dogodek (literal event) dogodka X je eden izmed X ali $\overline{X}$, kjer $\overline{X} = 1 - X$ označuje komplement od X. Če je $\mathbf{C}$ končna množica dogodkov, potem definiramo *množico dejanskih dogodkov* množice $\mathbf{C}$ kot množico elementov množice $\mathbf{C}$ in njihovih komplementov:

$$\mathbb{L}(\mathbf{C}) \equiv \mathbf{C} \cup \{\overline{X} \mid X \in \mathbf{C}\}.$$

Vrednost dogodka $L \in \mathbb{L}(\mathbf{C})$ pri vrednosti $\mathbf{c}$ od $\mathbf{C}$ označimo z $(L)_{\mathbf{c}}$ ali $(L)_{\mathbf{C}=\mathbf{c}}$. Vrednost $(L)_{\mathbf{c}}$ je iz množice $\{0, 1\}$.

Konjunkcijo množice dejanskih dogodkov $\mathbf{B} = \{F_1, ..., F_m\} \subseteq \mathbb{L}(\mathbf{C})$ definiramo kot

$$\bigwedge(\mathbf{B}) \equiv \prod_{i=1}^{m} F_i = \min\{F_1, ..., F_m\}.$$

Torej je konjunkcija $\bigwedge(\mathbf{B})$ slučajna spremenljivka, ki pri osebku ω zavzame vrednost $\prod_{i=1}^{m} F_i(\omega) = \min\{F_1(\omega), ..., F_m(\omega)\}$. Velja $\bigwedge(\mathbf{B}) = 1$, če in samo če $F_i = 1$ za vse i. Definiramo $B_1 \wedge B_2 \equiv \bigwedge\{B_1, B_2\}$. Naj bo $\mathbf{C} = \{X_1, \ldots, X_s\}$ ter $\mathbf{c} = (x_1, \ldots, x_s)$. Označimo z $(\bigwedge(\mathbf{B}))_{\mathbf{c}}$ oziroma $(\bigwedge(\mathbf{B}))_{\mathbf{C}=\mathbf{c}}$ vrednost konjunkcije množice $\mathbf{B}$, v primeru, da $\mathbf{C}$ zavzame vrednost $\mathbf{c}$. Vrednost $(\bigwedge(\mathbf{B}))_{\mathbf{c}}$ je iz množice $\{0, 1\}$.

3.2 Zadosten vzrok

Definicija 3.1 *(Zadosten vzrok). Naj bo $\Omega^* \subseteq \Omega$, $\Omega^* \neq \emptyset$ in naj bo $\mathbb{L}(\boldsymbol{C})$ množica dejanskih dogodkov. Zadosten vzrok za D (glede na $\mathbf{C}$) v podpodulaciji Ω^* je podmnožica $\boldsymbol{B} \subset \mathbb{L}(\boldsymbol{C})$, ki ima naslednjo lastnost: pri vrednosti $\boldsymbol{c} \in \{0,1\}^{|\mathbf{C}|}$, taki da $(\bigwedge(\boldsymbol{B}))_{\boldsymbol{c}} = 1$, velja $D_{\boldsymbol{c}}(\omega) = 1$ za vse $\omega \in \Omega^*$. Pri tem zahtevamo obstoj vsaj enega $\boldsymbol{c}^*$, da $(\bigwedge(\boldsymbol{B}))_{\boldsymbol{c}^*=1}$.*

Definicija zadostnega vzroka za ω^* zahteva veljavnost pogoja

$$D_{\mathbf{B}=\mathbf{1},\ \mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 1 \tag{3.1}$$

pri vsaki vrednosti $\mathbf{c}_2^*$.

Primer 3.3 *Upoštevajmo spet primer 3.2 o zadostnem vzroku tuberkuloze. Množica dogodkov $\boldsymbol{C}$ je enaka naslednji množici:*

$$\boldsymbol{C} = \{X_1, X_2, X_3, X_4, X_5, X_6\},$$

kjer spremenljivka X_1 označuje dobro ventilacijo, oziroma $X_1(\omega) = 1$, če osebek ω živi v prostoru z dobro ventilacijo, $X_1(\omega) = 0$ sicer, podobno X_2 označuje zadostni življenjski prostor, X_3 označuje dobro prehrano, X_4 označuje izpostavljenost tuberkulozi, X_5 označuje cepljenje proti tuberkulozi in X_6 označuje obolenje z AIDSom. Množica dejanskih dogodkov množice $\boldsymbol{C}$ je enaka

$$\mathbb{L}(\boldsymbol{C}) = \{X_1, \overline{X}_1, X_2, \overline{X}_2, X_3, \overline{X}_3, X_4, \overline{X}_4, X_5, \overline{X}_5\}.$$

Spremenljivka D označuje obolenje s tuberkulozo, oziroma $D(\omega) = 1$, če ima osebek ω tuberkulozo, $D(\omega) = 0$ sicer. Zadosten vzrok prvega osebka ω_1 iz primera 3.2 je $\boldsymbol{B} = \{B_1, B_2, B_3, B_4, B_5\}$, kjer $B_1 = \overline{X}_1$, $B_2 = \overline{X}_2$, $B_3 = \overline{X}_3$, $B_4 = X_4$ in $B_5 = \overline{X}_5$. Iz $(\bigwedge(\boldsymbol{B}))_{\boldsymbol{c}} = 1$ sledi $D_{\boldsymbol{c}}(\omega_1) = 1$, če $\boldsymbol{c} = (0, 0, 0, 1, 0, c_6)$, kjer je c_6 poljubna vrednost iz $\{0, 1\}$. Namreč veljajo naslednje enakosti:

$$\begin{aligned}
(\bigwedge(\boldsymbol{B}))_{\boldsymbol{C}=\boldsymbol{c}} &= (\bigwedge(\boldsymbol{B}))_{X_1=0, X_2=0, X_3=0, X_4=1, X_5=0, X_6=c_6} \\
&= (B_1 \wedge B_2 \wedge B_3 \wedge B_4 \wedge B_5)_{X_1=0, X_2=0, X_3=0, X_4=1, X_5=0, X_6=c_6} \\
&= (\overline{X}_1 \wedge \overline{X}_2 \wedge \overline{X}_3 \wedge X_4 \wedge \overline{X}_5)_{X_1=0, X_2=0, X_3=0, X_4=1, X_5=0, X_6=c_6} \\
&= (1 \wedge 1 \wedge 1 \wedge 1 \wedge 1) \\
&= 1.
\end{aligned} \tag{3.2}$$

Zadosten vzrok drugega osebka ω_2 iz primera 3.2 je le podmnožica prvega in je enak množici $\boldsymbol{B}' = \{B_1, B_2, B_4, B_5\}$. Iz $(\bigwedge(\boldsymbol{B}'))_{\boldsymbol{c}} = 1$ sledi $D_{\boldsymbol{c}}(\omega_2) = 1$, če $\boldsymbol{c} = (0, 0, c_3, 1, 0, c_6)$, kjer sta c_3 in c_6 poljubni vrednosti iz $\{0, 1\}$.

Poglejmo še zadosten vzrok tretjega osebka ω_3 iz primera 3.2: $\boldsymbol{B}'' = \{B_3, B_4, B_6\}$, kjer $B_6 = X_6$. Iz $(\bigwedge(\boldsymbol{B}''))_{\boldsymbol{c}} = 1$ sledi $D_{\boldsymbol{c}}(\omega_3) = 1$, če $\boldsymbol{c} = (c_1, c_2, 0, 1, c_5, 1)$, kjer so c_1, c_2 in c_5 poljubne vrednosti iz $\{0, 1\}$.

Definicijo zadostnega vzroka omejimo na neprazne množice Ω^* in tako preprečimo, da je vsaka množica $\mathbf{B}$ zadostni vzrok v prazni podpopulaciji. V definiciji zahtevamo obstoj vrednosti $\mathbf{c}^*$, da

$(\bigwedge(\mathbf{B}))_{\mathbf{c}^*} = 1$, z namenom, da se izognemo logično neskladnim konjunkcijam kot so npr. $X_1 \wedge \overline{X}_1$. Take konjunkcije v prazno zadovoljujejo pogoj iz definicije zadostnega vzroka.

Zadosten vzrok $\mathbf{B}$ ne more vsebovati bodisi spremenljivko X bodisi njen komplement $\overline{X}$. Definirajmo množico tistih podmnožic množice $\mathbb{L}(\mathbf{C})$, ki ne vsebujejo hkrati množice $X \in \mathbf{C}$ in množice $\overline{X}$:

$$\dot{\mathbb{P}}(\mathbb{L}(\mathbf{C})) \equiv \{\mathbf{B} \mid \mathbf{B} \subset \mathbb{L}(\mathbf{C}), \{X, \overline{X}\} \not\subseteq \mathbf{B}, \forall X \in \mathbf{C}\}.$$

Če $\mathbf{B} \in \dot{\mathbb{P}}(\mathbb{L}(\mathbf{C}))$, $|\mathbf{B}| = |\mathbf{C}|$ ter za vsak $C \in \mathbf{C}$ je ena sama množica izmed C in $\overline{C}$ vsebovana v $\mathbf{B}$, potem neka vrednost $\mathbf{b}$ množice $\mathbf{B}$ inducira eno samo vrednost $\mathbf{c}$ množice $\mathbf{C}$ ter obratno.

S formulacijo sledečih trditev poudarimo, da se zadostnost glede na neko množico $\mathbf{C}$ prenaša na podmnožice $\mathbf{C}$-ja, če te vsebujejo zadosten vzrok, ter, da se zadostnost v neki podpopulaciji Ω^* prenaša na manjše podpopulacije v Ω^*.

Trditev 3.2 *Če je $\boldsymbol{B}$ zadosten vzrok za D glede na $\boldsymbol{C}$ v Ω^*, potem je $\boldsymbol{B}$ zadosten vzrok za D glede na vsako množico $\boldsymbol{C}^*$ v Ω^*, kjer $\boldsymbol{B} \subseteq \boldsymbol{C}^* \subseteq \boldsymbol{C}$.*

Primer 3.4 *Množica $\boldsymbol{B}$ iz primera 3.3 je zadosten vzrok glede na $\boldsymbol{C}$ in tudi glede na $\boldsymbol{C}^* = \{X_1, X_2, X_3, X_4, X_5\}$, kjer $\boldsymbol{B} \subseteq \boldsymbol{C}^* \subseteq \boldsymbol{C}$.*

Zadostnost glede na neko množico $\mathbf{C}$ se ne prenaša na večje množice, ki vsebujejo $\mathbf{C}$: če je množica $\mathbf{B}$ zadosten vzrok za D glede na $\mathbf{C}$ v Ω^*, ni nujno tudi zadosten vzrok glede na neko večjo množico $\mathbf{C}' \supset \mathbf{C}$. Namreč lahko dogodek $X_7 \in \mathbf{C}' \setminus \mathbf{C}$ vpliva na dogodke iz zadostnega vzroka $\mathbf{B}$.

Trditev 3.3 *Če je $\boldsymbol{B}$ zadosten vzrok za D glede na $\boldsymbol{C}$ v Ω^*, potem je $\boldsymbol{B}$ zadosten vzrok za D glede na $\boldsymbol{C}$ v vsaki podmnožici $\Omega^{**} \subseteq \Omega^*$, kjer $\Omega^{**} \neq \emptyset$.*

Opomba 3.1 *Iz definicije zadostnega vzroka, sledi: če je $\boldsymbol{B}$ zadosten vzrok za D glede na $\boldsymbol{C}$, potem je vsaka množica $\boldsymbol{A} \subset \mathbb{L}(\boldsymbol{C})$, ki vsebuje $\boldsymbol{B}$, zadosten vzrok za D glede na $\boldsymbol{C}$.*

3.3 Minimalen zadosten vzrok

Idealen zadosten vzrok vsebuje le komponente, ki so nujno potrebne, da iz vzroka sledi posledica. Točneje se idealen zadosten vzrok izogne komponentam, brez katerih se izid vseeno pripeti, torej niso odločilnega pomena. Tak zadosten vzrok imenujemo minimalen zadosten vzrok.

Definicija 3.4 *(Minimalen zadosten vzrok). Minimalen zadosten vzrok za D (glede na $\boldsymbol{C}$) v podpopulaciji Ω^* je množica $\boldsymbol{B} \subset \mathbb{L}(\boldsymbol{C})$, ki je zadosten vzrok za D v Ω^* in taka, da nobena prava podmnožica $\boldsymbol{B}^* \subset \boldsymbol{B}$ ni zadosten vzrok za D v Ω^*.*

Definicija minimalnega zadostnega vzroka za ω^* zahteva za vsako množico $L \in \mathbf{B}$ obstoj vrednosti $\mathbf{c}_2^L$, da velja

$$D_{\mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0,\ \mathbf{C}_2=\mathbf{c}_2^L}(\omega^*) = 0. \tag{3.3}$$

Opomba 3.2 *Minimalnost zadostnega vzroka se ne prenaša na podmnožice: če je $\boldsymbol{B}$ minimalen zadosten vzrok za D glede na $\boldsymbol{C}$ v Ω^*, potem ni minimalen zadosten vzrok glede na $\boldsymbol{C}^* \subset \boldsymbol{C}$. Poglejmo protiprimer.*

Primer 3.5 *Poglejmo primer 2.1 in tabelo 1. Opazimo, da pri drugem osebku ω_2 je množica $\{X_1, X_3\}$ minimalen zadosten vzrok glede na $\{X_1, X_2, X_3\}$. Spoznali pa bomo, da množica $\{X_1, X_3\}$*

vezana na osebek ω_2 ni minimalen zadosten vzrok glede na $\{X_1, X_3\}$ pri predpostavki, da dogodek X_2 ni posledica dogodkov X_1 in X_3. Ta predpostavka zahteva, da za vse x_1, x_3 velja

$$(X_2)_{X_1=x_1,X_3=x_3}(\omega_2) = X_2(\omega_2).$$

Minimalnost v tem primeru ne velja, ker je X_3 zadosten vzrok za D glede na $\{X_1, X_3\}$. To opazimo iz naslednjih enačb, katere dobimo pri vrednosti $X_2(\omega_2) = 0$:

$$D_{X_1=0,X_3=1}(\omega_2) = D_{X_1=0,X_2=0,X_3=1(\omega_2)} = 1,$$

$$D_{X_1=1,X_3=1}(\omega_2) = D_{X_1=1,X_2=0,X_3=1(\omega_2)} = 1.$$

Opomba 3.3 *Minimalnost zadostnega vzroka v neki podpopulaciji Ω^* se ne prenaša na manjše podpopulacije v Ω^*: v primeru, da je $\boldsymbol{B}$ minimalen zadosten vzrok za D glede na $\boldsymbol{C}$ v Ω^*, potem $\boldsymbol{B}$ ni nujno minimalen zadosten vzrok za D glede na $\boldsymbol{C}$ v $\Omega^{**} \subseteq \Omega^*$. V posebnem primeru lahko velja, da množica $\boldsymbol{B}$ ni minimalen zadosten vzrok za D v $\{\omega\}$ za noben $\omega \in \Omega^*$.*

3.4 Minimalen zadosten vzrok pri pogoju monotonosti

V naslednjem poglavju bomo obravnavali zadosten pogoj minimalnosti v primeru, da imajo nekatere komponente zadostnega vzroka pozitiven monoton učinek na izid. Definirajmo najprej pozitiven monoton učinek, s pomočjo katerega bomo sestavili zadosten pogoj minimalnosti. Dogodek pozitivno monotono učinkuje na izid, če prispeva k uresničitvi izida ali pa ne vpliva na izid: izida ne prepreči.

Definicija 3.5 *(Pozitiven monoton učinek). Dogodek $B_i \in \boldsymbol{B}$ pozitivno monotono učinkuje na D glede na množico $\boldsymbol{B}$ v populaciji Ω, če za vse $\omega \in \Omega$ in vrednosti $\boldsymbol{b}$ spremenljivk iz $\boldsymbol{B} \setminus B_i$ velja*

$$D_{\boldsymbol{B}\setminus\{B_i\}=\boldsymbol{b},B_i=1}(\omega) \geq D_{\boldsymbol{B}\setminus\{B_i\}=\boldsymbol{b},B_i=0}(\omega).$$

Dogodek L *negativno monotono učinkuje* glede na $\mathbf{B} \cup \{L\}$, če $\overline{L}$ pozitivno monotono učinkuje glede na $\mathbf{B} \cup \{\overline{L}\}$. *Množica dogodkov $\boldsymbol{A}$ pozitivno monotono učinkuje* na D glede na množico $\mathbf{B}$ v populaciji Ω, če vsak $B_i \in \mathbf{A}$ pozitivno monotono učinkuje na D glede na množico $\mathbf{B}$ v populaciji Ω.

V primeru, da postavimo v zgornjo neenačbo enačaj za vse ω, potem B_i nima učinka na D glede na $\mathbf{B}$. Poudarimo, da neenakost iz definicije velja pri vsakem osebku iz podpopulacije, zato je pogoj zelo strog.

Pogoj monotonega učinka je izpolnjen v številnih konkretnih primerih in nam olajša preverjanje pogojev minimalnosti. Pomembno je, da se zavedamo prisotnosti pozitivnega učinka. Poglejmo trditev o zadostnem pogoju minimalnosti. Naj $A\dot{\cup}B$ označuje disjunktno unijo množic A in B.

Trditev 3.6 *Naj bo $\boldsymbol{C} = \boldsymbol{C}_1\dot{\cup}\boldsymbol{C}_2$, $\boldsymbol{B} \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}_1))$, $|\boldsymbol{B}| = |\boldsymbol{C}_1|$. Naj $\boldsymbol{C}_2$ pozitivno monotono učinkuje na D glede na $\boldsymbol{C}$. Če veljata pogoja*

$$\begin{aligned} &(i)\ D_{\boldsymbol{B}=\boldsymbol{1},\boldsymbol{C}_2=\boldsymbol{0}}(\omega^*) = 1, \\ &(ii)\ D_{\boldsymbol{B}\setminus\{L\}=\boldsymbol{1},L=0,\boldsymbol{C}_2=0}(\omega^*) = 0 \quad \text{za vse } L \in \boldsymbol{B}, \end{aligned} \tag{3.4}$$

potem je množica $\boldsymbol{B}$ minimalen zadosten vzrok za D glede na $\boldsymbol{C}$ pri ω^*.

Pogoja iz trditve sta šibkejša od pogoja iz definicije minimalnega zadostnega vzroka. Namreč, če ima $\mathbf{C}_2$ pozitiven monoton učinek na D, potem zadošča preveriti pogoj $D_{\mathbf{B=1,C_2=c_2}}(\omega^*) = 1$ le pri vrednosti $\mathbf{c}_2 = \mathbf{0}$.

Dokaz: Iz predpostavke monotonosti in pogoja (i) sledi

$$D_{\mathbf{B=1},\mathbf{C}_2=\mathbf{c}_2}(\omega^*) \geq D_{\mathbf{B=1},\mathbf{C}_2=\mathbf{0}}(\omega^*) = 1$$

za vsak $\mathbf{c}_2$. Torej je $\mathbf{B}$ zadosten vzrok za D glede na $\mathbf{C}$ za ω^*. Minimalnost sledi iz pogoja (ii), saj definicija minimalnega zadostnega vzroka (za ω^*) zahteva pri vsaki množici $L \in \mathbf{B}$ obstoj vrednosti $\mathbf{c}_2^L$, da velja $D_{\mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0,\ \mathbf{C}_2=\mathbf{c}_2^L}(\omega^*) = 0$. □

3.5 Odločilna in neodvečna množica zadostnih vzrokov

V sledečem poglavju bomo obravnavali dve vrsti množic zadostnih vzrokov oziroma dve vrsti množic $\mathfrak{B} = \{\mathbf{B}_1, \ldots, \mathbf{B}_n\} \subseteq \dot{\mathbb{P}}(\mathbb{L}(\mathbf{C}))$, kjer je vsaka množica $\mathbf{B}_i$ zadosten vzrok.

Množica zadostnih vzrokov $\mathfrak{B}$ je *odločilna* (determinative), v primeru, da ima naslednjo lastnost: če se dogodek D pripeti, potem se je nujno pripetil eden izmed zadostnih vzrokov iz množice $\mathfrak{B}$. Pojem odločilnosti ne zahteva samo, da množica zadostnih vzrokov $\mathfrak{B}$ povzroči posledico D, ampak tudi, da posledica sledi nujno iz množice zadostnih vzrokov $\mathfrak{B}$.

Ko je vsak izmed zadostnih vzrokov v množici $\mathfrak{B}$ potreben, da je množica odločilna, potem pravimo, da je množica $\mathfrak{B}$ *neodvečna* (nonredundant). Spomnimo se, da so zadostni vzroki $\mathbf{B}_i$ iz $\dot{\mathbb{P}}(\mathbb{L}(\mathbf{C}))$, saj ne vsebujejo hkrati množice $X \in \mathbf{C}$ in množice $\overline{X}$.

Pred formalno definicijo odločilnih množic poglejmo še definicijo disjunkcije slučajnih spremenljivk:

$$\bigvee(\{Z_1, \ldots, Z_p\}) \equiv \max\{Z_1, \ldots, Z_p\}.$$

Torej je disjunkcija $\bigvee(\{Z_1, \ldots, Z_p\})$ slučajna spremenljivka, ki pri osebku ω zavzame vrednost $\max\{Z_1(\omega), ..., Z_m(\omega)\}$. Velja $\bigvee(\{Z_1, \ldots, Z_p\}) = 1$, če in samo če velja $Z_j = 1$ za nek j. Definirajmo $Z_1 \vee Z_2 \equiv \bigvee\{Z_1, Z_2\}$. Če je $\mathfrak{B} = \{\mathbf{B}_1, \ldots, \mathbf{B}_q\}$ množica množic dejanskih dogodkov, potem definiramo

$$\bigwedge\bigvee(\mathfrak{B}) \equiv \bigwedge_i \left(\bigvee(\mathbf{B}_i)\right).$$

Definicija 3.7 *(Odločilna množica zadostnih vzrokov). Množica zadostnih vzrokov za D, $\mathfrak{B} = \{\boldsymbol{B}_1, \ldots, \boldsymbol{B}_n\} \subseteq \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}))$, je odločilna za D (glede na $\boldsymbol{C}$) v podpopulaciji Ω^*, če za vsak $\omega \in \Omega^*$ in za vse $\boldsymbol{c}$ velja*

$$D_{\boldsymbol{c}}(\omega) = 1 \Leftrightarrow (\bigvee\bigwedge(\mathfrak{B}))_{\boldsymbol{c}} = 1.$$

Odločilno množico zadostnih vzrokov za D imenujemo *zadostni kavzalni model.* Vsaka podpopulacija Ω^*, za katero obstaja odločilna množica zadostnih vzrokov, ima enak vektor možnih izidov za D tj. $\mathcal{D}(\mathbf{C}, \omega) = \mathcal{D}(\mathbf{C}, \omega')$ za vse $\omega,\ \omega' \in \Omega^*$. Namreč za $\omega,\ \omega' \in \Omega^*$ in $\mathbf{c}$ velja $D_{\mathbf{c}}(\omega) = 1$, če in samo če $(\bigvee\bigwedge(\mathfrak{B}))_{\mathbf{c}} = 1$, kar velja, če in samo če $D_{\mathbf{c}}(\omega') = 1$.

Definicija 3.8 *(Neodvečna množica zadostnih vzrokov). Odločilna množica zadostnih vzrokov $\mathfrak{B}$ za D je neodvečna (v Ω^* glede na $\boldsymbol{C}$), če ne obstaja prava podmnožica $\mathfrak{B}^* \subset \mathfrak{B}$, ki je odločilna za D.*

Primerjajmo definicijo minimalnega zadostnega vzroka ter neodvečne odločilne množice zadostnih vzrokov. Obe zahtevata odsotnost pravih podmnožic. Osnovna razlika je v tem, da prva se nanaša na konjunkcije, druga pa na disjunkcije. Namreč pri minimalnem zadostnem vzroku $\mathbf{B}$ je vsak element konjunkcije $\bigwedge(\mathbf{B})$ potreben, da sledi posledica D. Pri neodvečnih odločilnih množicah zadostnih vzrokov $\mathfrak{B}$ pa je vsak element disjunkcije $(\bigvee\bigwedge(\mathfrak{B}))$ potreben, da je množica $\mathfrak{B}$ odločilna.

Primer 3.6 *Obravnavajmo dve neodvečni množici, ki sta vezani na drugega osebka iz primera 2.1:*

$$\mathfrak{B}_1 = \{\{X_1, X_2\}, \{X_2, \overline{X_3}\}, \{\overline{X}_2, X_3\}\},$$

$$\mathfrak{B}_2 = \{\{X_1, X_3\}, \{X_2, \overline{X_3}\}, \{\overline{X}_2, X_3\}\}.$$

Poglejmo najprej, zakaj je $\mathfrak{B}_1$ *neodvečna. Pogoj* $D_{\boldsymbol{C}=\boldsymbol{c}}(\omega_2) = 1$ *velja, če in samo če množica* $\boldsymbol{C} = \{X_1, X_2, X_3\}$ *zavzame vrednosti* $(c_1, 0, 1)$, $(c_2, 1, 0)$ *ali* $(1, c_3, 1)$ *pri poljubnih* c_1, c_2, c_3 *iz* $\{0, 1\}$. *To pomeni, da* $\overline{X_2} \wedge X_3 = 1$, $X_2 \wedge \overline{X}_3 = 1$ *ali* $X_1 \wedge X_2 = 1$. *Torej* $((X_1 \wedge X_2) \vee (X_2 \wedge \overline{X}_3) \vee (\overline{X_2} \wedge X_3))_{\boldsymbol{c}} = 1$ *in* $\mathfrak{B}_1$ *je odločilna množica. Vsaka podmnožica iz množice* $\mathfrak{B}_1$ *je potrebna, da se izpolni pogoj odločilnosti, torej je* $\mathfrak{B}_1$ *tudi neodvečna množica zadostnih vzrokov za drugi osebek* ω_2.

Podobno vidimo, da je tudi množica $\mathfrak{B}_2$ *neodvečna, saj* $D_{\boldsymbol{c}}(\omega_2) = ((X_1 \wedge X_3) \vee (X_2 \wedge \overline{X}_3) \vee (\overline{X_2} \wedge X_3))_{\boldsymbol{c}}$.

3.6 Reprezentacije zadostnega vzroka

Pogoj iz definicije zadostnega vzroka je težko izpolnjen pri vsakem osebku v populaciji: običajno imajo različne podpopulacije različne zadostne vzroke. Definirajmo množico zadostnih vzrokov različnih podpopulaciji s pomočjo *reprezentacije zadostnega vzroka*. To je množica populacijskih zadostnih vzrokov in binarnih slučajnih spremenljivk, ki označujejo posamezne podpopulacije.

Definicija 3.9 *(Reprezentacija zadostnega vzroka). Reprezentacija zadostnega vzroka* $(\boldsymbol{A}, \mathfrak{B})$ *za* $\mathcal{D}(\boldsymbol{C}; \Omega)$ *je urejena množica slučajnih spremenljivk* $\boldsymbol{A} = \langle A_1, \ldots, A_p \rangle$, *ki so definirane za vse osebke v populaciji, zavzemajo vrednosti iz* $\{0, 1\}$ *ter* $(A_i)_{\boldsymbol{c}} = A_i$ *za vse* i, $\boldsymbol{c}$, *in množica* $\mathfrak{B} = \langle \boldsymbol{B}_1, \ldots, \boldsymbol{B}_p \rangle$, *kjer je* $\boldsymbol{B}_i \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}))$ *taka, da za vse* ω, $\boldsymbol{c}$ *velja:*

$$D_{\boldsymbol{c}}(\omega) = 1 \Leftrightarrow \text{ za nek } j \text{ velja } A_j(\omega) = 1 \text{ in } (\bigwedge(\boldsymbol{B}_i))_{\boldsymbol{c}} = 1.$$

Spremenljivke $A_i \in \mathbf{A}$ označujejo podpopulacije: če je osebek ω v i-ti podpopulaciji, potem velja $A_i(\omega) = 1$.

Iz leve smeri zgornjega pogoja vidimo, da je $\mathbf{B}_j$ zadosten vzrok v j-ti podpopulaciji: če za nek j velja $A_j(\omega) = 1$ potem $D_{\mathbf{B}_j=\mathbf{1}}(\omega) = 1$. Iz desne smeri zgornjega pogoja pa ugotovimo naslednjo lastnost: če se je pripetil izid pri osebku ω, $D_{\mathbf{c}}(\omega) = 1$, potem pripada osebek ω j-ti podpopulaciji, ki ima zadosten vzrok enak $\mathbf{B}_j \subset \mathbf{B}^{[\mathbf{c}]}$.

Slučajne spremenljivke A_i in množice $\mathbf{B}_i$ naravno nastopajo v parih, ki jih določa ureditev $\mathbf{A}$ in $\mathfrak{B}$. Pravimo, da je par $(A_i, \mathbf{B}_i)$ *prisoten* v reprezentaciji. Če je $\mathbf{C} = \{X_1, \ldots, X_s\}$ in $\mathbf{c} = \{x_1, \ldots, x_s\}$, potem iz zahteve $(A_i)_{\mathbf{c}} = A_i$ sledi $(A_i)_{X_1=x_1,\ldots X_s=x_s} = A_i$ in torej posegi na X_i ne vplivajo na A_i. Sledi neenakost $A_i \neq X_j$ za vsak i ter za vsak j. Obenem $\mathbf{A} \cap \mathbf{C} = \emptyset$. Pogoj $(A_i)_{\mathbf{c}} = A_i$ nam omogoča, da gledamo na spremenljivko A_i kot na označevalko podpopulacije i. Namreč populacije so vnaprej določene in neodvisne od vrednosti $\mathbf{c}$ od $\mathbf{C}$. Iz zgornje definicije neposredno sledi naslednja trditev.

Trditev 3.10 *Če* $(\boldsymbol{A}, \mathfrak{B})$ *je reprezentacija zadostnega vzroka za* $\mathcal{D}(\boldsymbol{C}; \Omega)$, $\boldsymbol{A} = \langle A_1, \ldots, A_p \rangle$ *ter* $\mathfrak{B} = \langle \boldsymbol{B}_1, \ldots, \boldsymbol{B}_p \rangle$, *potem je* $\boldsymbol{B}_i$ *zadostni vzrok za* D *v podpopulaciji, kjer velja* $A_i(\omega) = 1$.

Definicija množice $\mathfrak{B}$ iz reprezentacije zadostnega vzroka ter definicija odločilne množice zadostnih vzrokov sta podobni, vendar se konceptualno razlikujeta. Poglejmo natančneje podobnosti in razlike. Pogoj

$$D_{\mathbf{c}}(\omega) = 1 \Rightarrow (\bigvee \bigwedge(\mathfrak{B}))_{\mathbf{c}} = 1$$

velja v obeh definicijah. Obraten pogoj pa velja le v definiciji odločilne množice. Če upoštevamo poljuben osebek ω in poljubno množico zadostnih vzrokov $\mathbf{B}_j$ iz reprezentacije, za katero velja $(\bigwedge(\mathbf{B}_i))_{\mathbf{c}} = 1$, potem ni nujno, da je osebek ω v j-ti populaciji, oziroma ne velja nujno $A_j(\omega) = 1$ in posledično ne moremo trditi, da velja $D_{\mathbf{c}}(\omega) = 1$.

Torej $\mathfrak{B}$ iz reprezentacije zadostnega vzroka ni odločilna množica za celotno populacijo; vsekakor pa obstaja podmnožica $\mathfrak{B}$-ja, ki je odločilna za manjšo podpopulacijo. V naslednji trditvi bomo to podmnožico opisali.

Trditev 3.11 *Naj bo $(\boldsymbol{A}, \mathfrak{B})$ reprezentacija zadostnega vzroka za $\mathcal{D}(\boldsymbol{C}; \Omega)$ in naj bo $\boldsymbol{A}^* \subseteq \boldsymbol{A}$, taka da $\emptyset \neq \Omega^{\boldsymbol{A}^*}_{\boldsymbol{A} \setminus \boldsymbol{A}^*} \equiv \{\omega \in \Omega \mid$ za vse $A_i \in \boldsymbol{A}, A_i(\omega) = 1 \Leftrightarrow A_i \in \boldsymbol{A}^*\}$. Potem je množica*

$$\mathfrak{B}^{\boldsymbol{A}^*} \equiv \{\boldsymbol{B}_i \mid \boldsymbol{B}_i \in \mathfrak{B}; A_i \in \boldsymbol{A}^*\}$$

odločilna množica zadostnih vzrokov (glede na $\boldsymbol{C}$) v $\Omega^{\boldsymbol{A}^*}_{\boldsymbol{A} \setminus \boldsymbol{A}^*}$.

Podpopulacija $\Omega^{\mathbf{A}^*}_{\mathbf{A} \setminus \mathbf{A}^*}$ je sestavljena iz osebkov ω, za katere velja $A_i(\omega) = 1$ za vse $A_i \in \mathbf{A}^*$ ter velja $A_j(\omega) = 0$ za vse $A_j \in \mathbf{A} \setminus \mathbf{A}^*$. Torej spremenljivke iz $\mathbf{A}^*$ karakterizirajo podpopulacijo $\Omega^{\mathbf{A}^*}_{\mathbf{A} \setminus \mathbf{A}^*}$. Množica $\mathfrak{B}^{\mathbf{A}^*}$ je pa sestavljena iz množic, ki nastopajo v paru s spremenljivkami iz $\mathbf{A}^*$.

Primer 3.7 *V posebnem primeru, kjer ni podpopulacij velja*

$$\Omega^{\boldsymbol{A}^*}_{\boldsymbol{A} \setminus \boldsymbol{A}^*} = \{\omega \in \Omega \mid A_1(\omega) = 1, \ldots, A_p(\omega) = 1\} = \Omega$$

ter $\boldsymbol{A}^* = \boldsymbol{A}$ in se definiciji reprezentacije zadostnega vzroka ter odločilne množice zadostnih vzrokov ujemata. Velja namreč $\mathfrak{B} = \mathfrak{B}^{\boldsymbol{A}^*}$.

Primer 3.8 *Poglejmo še primer, kjer je osebek ω^* v prvih dveh podpopulacijah, ki jih določata A_1 ter A_2, torej $A_1(\omega^*) = 1$ ter $A_2(\omega^*) = 1$. Velja $\boldsymbol{A}^* = \{A_1, A_2\}$, $\Omega^{\boldsymbol{A}^*}_{\boldsymbol{A} \setminus \boldsymbol{A}^*} = \{\omega^*\}$ in $\mathfrak{B}^{\boldsymbol{A}^*} = \{\boldsymbol{B}_1, \boldsymbol{B}_2\}$ je odločilna množica zadostnih vzrokov (glede na $\boldsymbol{C}$) v $\Omega^{\boldsymbol{A}^*}_{\boldsymbol{A} \setminus \boldsymbol{A}^*}$.*

Poglejmo dokaz trditve 3.11.

Dokaz: Naj bo $\omega \in \Omega^{\mathbf{A}^*}_{\mathbf{A} \setminus \mathbf{A}^*}$ in vrednost $\mathbf{c}$ poljubna. Pokazati moramo, da velja $D_{\mathbf{c}}(\omega) = 1$, če in samo če $(\bigvee \bigwedge(\mathfrak{B}^{\mathbf{A}^*}))_{\mathbf{c}} = 1$.

Predpostavimo najprej, da za ω velja $D_{\mathbf{c}}(\omega) = 1$. Iz definicije reprezentacije zadostnega vzroka sledi, da za ω obstaja j, tako da $A_j(\omega) = 1$ in $(\bigwedge(\mathbf{B}_j))_{\mathbf{c}} = 1$. Iz definicije množice $\Omega^{\mathbf{A}^*}_{\mathbf{A} \setminus \mathbf{A}^*}$ sledi $A_j \in \mathbf{A}^*$; iz definicije množice $\mathfrak{B}^{\mathbf{A}^*}$ pa $\mathbf{B}_j \in \mathfrak{B}^{\mathbf{A}^*}$. Torej $(\bigvee \bigwedge(\mathfrak{B}^{\mathbf{A}^*}))_{\mathbf{c}} \geq (\bigwedge(\mathbf{B}_j))_{\mathbf{c}} = 1$.

Predpostavimo še, da za nek j, $\mathbf{B}_j \in \mathfrak{B}^{\mathbf{A}^*}$, tako da $(\bigwedge(\mathbf{B}_j))_{\mathbf{c}} = 1$. Iz definicije množice $\mathfrak{B}^{\mathbf{A}^*}$ sledi $A_j \in \mathbf{A}^*$; iz definicije množice $\Omega^{\mathbf{A}^*}_{\mathbf{A} \setminus \mathbf{A}^*}$ pa sledi $A_j(\omega) = 1$. Nazadnje iz definicije reprezentacije zadostnega vzroka sledi $D_{\mathbf{c}}(\omega) = 1$. □

Sledi pomemben izrek, ki obravnava reprezentacije zadostnega vzroka in možne izide. Trdi, da za vsako množico možnih izidov $\mathcal{D}(\mathbf{C}; \Omega)$ lahko najdemo reprezentacijo zadostnega vzroka. V dokazu izreka bomo konstruirali reprezentacijo zadostnega vzroka s pomočjo *množice dejanskih dogodkov pri vrednosti $\boldsymbol{c}$ od $\boldsymbol{C}$*, katero definiramo

$$\mathbf{B}^{[\mathbf{c}]} \equiv \{L \mid L \in \mathbb{L}(\mathbf{C}), (L)_{\mathbf{c}} = \mathbf{1}\}.$$

Torej velja $|\mathbf{B}^{[\mathbf{c}]}| = |\mathbf{C}|$ in $\mathbf{B}^{[\mathbf{c}]} = 1$, če in samo če množica $\mathbf{C}$ zavzame vrednost $\mathbf{c}$. Množico $\mathbf{B}^{[\mathbf{c}]}$ sestavljajo torej množice iz $\mathbb{L}(\mathbf{C})$, ki zavzamejo vrednost 1, v primeru, da množica $\mathbf{C}$ zavzame vrednost $\mathbf{c}$. Natančneje, če je $\mathbf{C} = (Y_1, \ldots, Y_s)$ in je $\mathbf{c} = (c_1, \ldots, c_s)$, potem je $\mathbf{B}^{[\mathbf{c}]} = \{\tilde{Y}_1, \ldots, \tilde{Y}_s\}$, kjer

$$\tilde{Y}_i = \begin{cases} Y_i, & \text{če} \quad c_i = 1 \\ \overline{Y}_i, & \text{če} \quad c_i = 0. \end{cases}$$

Izrek 3.12 *Za vsak $\mathcal{D}(\boldsymbol{C}; \Omega)$ obstaja reprezentacija zadostnega vzroka $(\boldsymbol{A}, \mathfrak{B})$.*

Dokaz: Naj bo $p = 2^{|\mathbf{C}|}$. Definirajmo poljubno urejeno množico

$$\mathfrak{B} \equiv \{\mathbf{B} \mid \mathbf{B} \subseteq \dot{\mathbb{P}}(\mathbb{L}(\mathbf{C})),\ |\mathbf{B}| = |\mathbf{C}|\} \equiv \langle \mathbf{B}_1, \ldots, \mathbf{B}_p \rangle,$$

in $A_i(\omega) \equiv D_{\mathbf{B}_i = 1}(\omega)$. Množice $\mathbf{B}_i$ definirajmo na sledeči način. Naj bo $\mathbf{C} = (X_1, \ldots, X_s)$ ter naj bo $\mathbf{c} = (c_1, \ldots, c_s)$ poljuben vektor. Velja:

$$D_{\mathbf{C}=\mathbf{c}}(\omega) = 1 \quad \Leftrightarrow \quad D_{X_1 = c_1, \ldots, X_s = c_s}(\omega) = 1 \quad \Leftrightarrow \quad D_{\tilde{X}_1 = 1, \ldots, \tilde{X}_s = 1}(\omega) = 1,$$

kjer

$$\tilde{X}_i \equiv \begin{cases} X_i, & \text{če} \quad c_i = 1 \\ \overline{X}_i, & \text{če} \quad c_i = 0. \end{cases}$$

Definirajmo $\mathbf{B}_j \equiv \{\tilde{X}_1, \ldots, \tilde{X}_s\}$ oziroma $\mathbf{B}_j \equiv \mathbf{B}^{[\mathbf{c}]}$. Preverimo, da je definirana množica reprezentacija zadostnih vzrokov. Velja:

$$D_{\mathbf{c}}(\omega) = 1 \quad \Leftrightarrow \quad D_{\mathbf{B}_j = \mathbf{1}}(\omega) = 1 \quad \Leftrightarrow \quad A_j(\omega) = 1 \text{ in } (\bigwedge(\mathbf{B}_j))_{\mathbf{c}} = 1.$$

Pogoj $(\bigwedge(\mathbf{B}_j))_{\mathbf{c}} = 1$ velja, saj je $\mathbf{B}_j = \mathbf{B}^{[\mathbf{c}]}$. □

Poudarimo, da smo v zgornjem dokazu konstruirali reprezentacijo zadostnega vzroka $(\mathbf{A}, \mathfrak{B})$ za množico možnih izidov $\mathcal{D}(\mathbf{C}; \Omega)$, tako da smo definirali $A_i(\omega) \equiv D_{\mathbf{B}_i = 1}(\omega)$ in $\mathbf{B}_j \equiv \mathbf{B}^{[\mathbf{c}]}$.

Primer 3.9 *Poglejmo spet primer 2.1. Množice zadostnih vzrokov, ki reprezentirajo $\mathcal{D}(\boldsymbol{C}, \Omega)$ v tabeli 1, so naslednje oblike:*

$$\begin{aligned} \mathfrak{B} = & \langle \boldsymbol{B}_1, \ldots, \boldsymbol{B}_s \rangle \\ = & \langle \{X_1, X_2, X_3\}, \{X_1, X_2, \overline{X}_3\}, \{X_1, \overline{X}_2, X_3\}, \{X_1, \overline{X}_2, X_3\}, \\ & \{\overline{X}_1, X_2, X_3\}, \{\overline{X}_1, X_2, X_3\}, \{\overline{X}_1, \overline{X}_2, X_3\}, \{\overline{X}_1, \overline{X}_2, \overline{X}_3\} \rangle \end{aligned}$$

ter

$$A_1(\omega) = \begin{cases} 1, & \text{če } \omega = 2 \\ 0, & \text{če } \omega \neq 2, \end{cases}$$

$A_4 = A_5 = A_8 = 0$ in $A_2 = A_3 = A_6 = A_7 = 1$. Množice $\boldsymbol{B}_i$ in A_i poiščemo s pomočjo konstrukcije iz zgornjega dokaza: iz pogojev $D_{(c_1, c_2, c_3)}(\omega) = 1$ dobimo množice $\boldsymbol{B}_k = \boldsymbol{B}^{[\boldsymbol{c}]}$ ter $A_i(\omega) \equiv D_{\boldsymbol{B}_i = 1}(\omega)$.

4 Ireducibilne komponente

V sledečem poglavju bomo obravnavali komponente zadostnega vzroka, ki so prisotne v vsaki reprezentacij zadostnega vzroka. Imenujemo jih *ireducibilne* komponente. Ireducibilna komponenta ne povzroči nujno izida, ampak je le del zadostnega vzroka v reprezentaciji. Običajno ni enaka niti zadostnemu vzroku, niti minimalnemu zadostnemu vzroku.

V primeru tuberkuloze je izpostavljenost tuberkulozi ireducibilna komponenta. Namreč mora biti nujno prisotna v večini zadostnih vzrokov populacije, sama pa ni zadostna, da povzroči bolezen.

Definicija 4.1 *(Ireducibilna množica). Množica $\boldsymbol{B} \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}))$ je ireducibilna za $\mathcal{D}(\boldsymbol{C},\Omega)$, če v vsaki reprezentaciji $(\boldsymbol{A},\mathfrak{B})$ za $\mathcal{D}(\boldsymbol{C},\Omega)$, obstaja $\boldsymbol{B}_i \in \mathfrak{B}$, da $\boldsymbol{B} \subset \boldsymbol{B}_i$.*

V primeru 3.6 smo obravnavali dve neodvečni množici zadostnih vzrokov osebka ω_2. Obe množici sta vsebovali podmnožici $\{X_2, \overline{X}_3\}$ in $\{\overline{X}_2, X_3\}$. Zanima nas, ali sta ti dve množici ireducibilni za $\mathcal{D}(\mathbf{C},\Omega)$ v tabeli 1. To bomo z lahkoto ugotovili s pomočjo naslednjega izreka, v katerem bomo karakterizirali ireducibilne množice.

4.1 Zadosten in potreben pogoj ireducibilnosti

Izrek 4.2 *Naj bo $\boldsymbol{C} = \boldsymbol{C}_1 \dot{\cup} \boldsymbol{C}_2$, $\boldsymbol{B} \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}_1))$, $|\boldsymbol{B}| = |\boldsymbol{C}_1|$. Množica $\boldsymbol{B}$ je ireducibilna za $\mathcal{D}(\boldsymbol{B},\Omega)$, če in samo če obstajajo $\omega^* \in \Omega$ in vrednosti $\boldsymbol{c}_2^*$ za $\boldsymbol{C}_2$, da velja:*
(i) $D_{\boldsymbol{B}=\boldsymbol{1},\ \boldsymbol{C}_2=\boldsymbol{c}_2^*}(\omega^*) = 1$;
(ii) za vse $L \in \boldsymbol{B}$, $D_{\boldsymbol{B}\setminus\{L\}=\boldsymbol{1},\ L=0,\ \boldsymbol{C}_2=\boldsymbol{c}_2^*}(\omega^*) = 0$.

Ireducibilna množica je torej (i) del zadostnega vzroka (nekega osebka), ki (ii) je nujno potreben, da vzrok izpolnjuje definicijo zadostnega vzroka. Natančneje, množica $\mathbf{B}$ je ireducibilna, če obstaja vsaj en osebek ω^*, da veljata naslednja pogoja: (i) množica $\mathbf{B}$ je del zadostnega vzroka $\mathbf{E}$ za osebek ω^*; (ii) če se $\mathbf{B}$ ne pripeti v celoti, potem $\mathbf{E}$ ni več zadosten vzrok.

Pogoja (i) in (ii) sta ekvivalentna naslednjemu pogoju:

$$D_{\mathbf{B}=\mathbf{1},\ \mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) - \sum_{L\in\mathbf{B}} D_{\mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0,\ \mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) > 0. \tag{4.1}$$

V naslednjih poglavjih bomo uporabili zgornji pogoj pri empiričnem preverjanju ireducibilnosti. Dokažimo izrek 4.2.

Dokaz: ($\Rightarrow$) Prilagodimo dokaz izreka 3.12 o obstoju reprezentacije zadostnega vzroka. Predpostavimo ireducibilnost $\mathbf{B}$ in neveljavnost vsaj ene izmed predpostavk (i) in (ii) za vse $\omega \in \Omega$ in vrednosti $\mathbf{c}_2^*$ od $\mathbf{C}_2$. Pokazali bomo, da obstaja taka reprezentacija $(\mathbf{A},\mathfrak{B})$ za $\mathcal{D}(\mathbf{C},\Omega)$, da za vse $\mathbf{B}_i \in \mathfrak{B}$ velja $\mathbf{B} \not\subseteq \mathbf{B}_i$, kar je v protislovju z ireducibilnostjo $\mathbf{B}$. Definirajmo naslednji množici s poljubnimi ureditvami:

$$\mathfrak{B}^\dagger \equiv \langle \mathbf{B}_i^\dagger \rangle \equiv \{\mathbf{B}^* \mid \mathbf{B}^* \in \dot{\mathbb{P}}(\mathbb{L}(\mathbf{C})),\ |\mathbf{B}^*| = |\mathbf{C}|,\ \mathbf{B} \not\subseteq \mathbf{B}^*\},$$
$$\mathfrak{B}^\ddagger \equiv \langle \mathbf{B}_i^\ddagger \rangle \equiv \{\mathbf{B}^* \mid \mathbf{B}^* \in \dot{\mathbb{P}}(\mathbb{L}(\mathbf{C})),\ |\mathbf{B}^*| = |\mathbf{C}| - 1,\ \mathbf{B} \setminus \mathbf{B}^* = \{L\},\ L \in \mathbf{B}\}.$$

Torej je $\mathfrak{B}^\dagger$ sestavljena iz množic, ki vsebujejo $|\mathbf{C}|$ dejanskih dogodkov in ne vsebujejo množice $\mathbf{B}$; $\mathfrak{B}^\ddagger$ pa je sestavljena iz možic, ki vsebujejo $|\mathbf{C}| - 1$ dejanskih dogodkov in vsebujejo vse razen enega elementa iz $\mathbf{B}$. Definirajmo še spremenljivke, ki ustrezajo $\mathbf{B}_i^\dagger \in \mathfrak{B}^\dagger$,

$$A_i^\dagger(\omega) \equiv D_{\mathbf{B}_i^\dagger=1}(\omega),$$

ter spremenljivke, ki ustrezajo $\mathbf{B}_i^\ddagger \in \mathfrak{B}^\ddagger$,

$$A_i^\ddagger(\omega) \equiv D_{\mathbf{B}_i^\ddagger=1, L_i=0}(\omega) D_{\mathbf{B}_i^\ddagger=1, L_i=1}(\omega),$$

kjer $\{L_i\} \equiv \mathbf{B}\setminus\mathbf{B}_i^\ddagger$. Iskana reprezentacija je $(\mathbf{A},\mathfrak{B}) \equiv (\mathbf{A}^\dagger \cup \mathbf{A}^\ddagger, \mathfrak{B}^\dagger \cup \mathfrak{B}^\ddagger)$, kjer $\mathbf{A}^\dagger \equiv \langle A_i^\dagger \rangle$, $\mathbf{A}^\dagger \equiv \langle A_i^\ddagger \rangle$. Preverimo, da je to reprezentacija zadostnega vzroka.

Za nek ω in $\mathbf{c}$ obstaja par $(A_j, \mathbf{B}_j)$ v $(\mathbf{A}, \mathfrak{B})$, da $A_j(\omega) = 1$ in $(\bigwedge(\mathbf{B}_j))_{\mathbf{c}} = 1$. Dokazujemo, da $D_{\mathbf{c}}(\omega) = 1$. Če je $A_j \in \mathbf{A}^\dagger$ potem $D_{\mathbf{C}=\mathbf{c}}(\omega) = D_{\mathbf{B}_j=1} = A_j(\omega) = 1$. Če pa $A_j \in \mathbf{A}^\ddagger$ potem $D_{\mathbf{C}=\mathbf{c}}(\omega) = D_{\mathbf{B}_j=1, L_i=c} = 1$, kjer zadnja enakost sledi iz $1 = A_i(\omega) = D_{\mathbf{B}_i=1,L_i=0}(\omega) D_{\mathbf{B}_i=1,L_i=1}(\omega)$.

Poglejmo obratno smer. Naj za nek ω in $\mathbf{c}$ velja $D_{\mathbf{c}}(\omega) = 1$. Obravnavajmo dva možna primera.

Najprej poglejmo primer, ko je $(\bigwedge(\mathbf{B}))_{\mathbf{c}} = 0$. Potem velja $\mathbf{B} \nsubseteq \mathbf{B}^{[\mathbf{c}]}$ in vzamemo tisti $\mathbf{B}_j^\dagger$, da $\mathbf{B}_j^\dagger = \mathbf{B}^{[\mathbf{c}]}$. Pripadajoči $A_j^\dagger$ je tak, da $A_j^\dagger(\omega) = D_{\mathbf{B}_i^\dagger=1}(\omega) = D_{\mathbf{c}}(\omega) = 1$.

Če pa velja $(\bigwedge(\mathbf{B}))_{\mathbf{c}} = 1$, potem za $c = (x_1, \ldots, x_s)$ velja

$$(\bigwedge(\mathbf{B}))_{\mathbf{C}=\mathbf{c}} = (\bigwedge(\mathbf{B}))_{X_1=x_1,\ldots,X_s=x_s} = (\bigwedge(\mathbf{B}))_{\mathbf{C}_1=\mathbf{c}_1,\mathbf{C}_2=\mathbf{c}_2} = 1,$$

kjer je $(\mathbf{c}_1, \mathbf{c}_2)$ particija $\mathbf{c}$. Pogoj (i) velja pri $\mathbf{c}_2^* = \mathbf{c}_2$, saj $D_{\mathbf{B}=\mathbf{1},\mathbf{C}_2^*=\mathbf{c}_2}(\omega) = D_{\mathbf{c}}(\omega) = 1$. Torej pogoj (ii) ne velja: obstaja $L \in \mathbf{B}$, da $D_{\mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0,\ \mathbf{C}_2=\mathbf{c}_2}(\omega) = 1$. Izberemo takega $\mathbf{B}_j^\ddagger$ iz $\mathfrak{B}^\ddagger$, da velja $\mathbf{B}_j^\ddagger = \mathbf{B}^{[\mathbf{c}]}$. Torej velja $(\bigwedge(\mathbf{B}_j^\ddagger))_{\mathbf{c}} = 1$ in $A_j^\ddagger(\omega) = 1$, saj $D_{\mathbf{c}}(\omega) = D_{\mathbf{B}\setminus\{L\}=\mathbf{1},\ L=1,\ \mathbf{C}_2=\mathbf{c}_2}(\omega) = 1$.

($\Leftarrow$) Predpostavimo veljavnost pogojev (i) ter (ii) za ω^* ter $\mathbf{c}_2^*$ in predpostavimo, da $\mathbf{B}$ ni ireducibilna. Iz negacije definicije ireducibilnosti obstaja taka reprezentacija $(\mathbf{A}, \mathfrak{B})$, da za vse $\mathbf{B}_i \in \mathfrak{B}$ velja $\mathbf{B} \nsubseteq \mathbf{B}_i$. Iz pogoja (i) velja $D_{\mathbf{B}=\mathbf{1},\mathbf{c}_2^*}(\omega^*) = 1$. Iz definicije reprezentacije sledi obstoj para $(A_j, \mathbf{B}_j)$, da $A_j(\omega^*) = 1$ in $(\bigwedge(\mathbf{B}_j))_{\mathbf{B}=\mathbf{1},\mathbf{c}_2^*} = 1$ oziroma $\mathbf{B}_j \subseteq \mathbf{B} \cup \mathbf{B}^{[\mathbf{c}_2^*]}$. Iz $\mathbf{B} \nsubseteq \mathbf{B}_i$ sledi obstoj $L \in \mathbf{B} \setminus \mathbf{B}_j$, ki ne vpliva na realizacijo $\mathbf{B}_j$ oziroma $(\bigwedge(\mathbf{B}_j))_{\mathbf{B}\setminus L=\mathbf{1},L=0,\mathbf{C}_2=\mathbf{c}_2^*} = 1$. Iz definicije reprezentacije ter $A_j(\omega^*) = 1$ in $(\bigwedge(\mathbf{B}_j))_{\mathbf{B}\setminus L=\mathbf{1},L=0,\mathbf{C}_2=\mathbf{c}_2^*} = 1$ sledi $D_{\mathbf{B}\setminus L=\mathbf{1},L=0,\mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 1$, kar je v protislovju s pogojem (ii). $\square$

Primer 4.1 *Poglejmo ali sta množici* $\{X_2, \overline{X}_3\}$ *in* $\{\overline{X}_2, X_3\}$ *iz primera 3.6 ireducibilni komponenti neodvečnih množic* $\mathfrak{B}_1$ *in* $\mathfrak{B}_2$ *pri osebku* ω_2. *Pri vrednosti* $c_2^* = 0$ *od* $C_2 = X_1$ *veljata naslednja dva pogoja:*

(i) $\boldsymbol{D}_{X_2=1,\overline{X}_3=1,C_2=c_2^*}(\omega_2) = \boldsymbol{D}_{010} = 1,$

(ii) $\boldsymbol{D}_{X_2=1,\overline{X}_3=0,C_2=c_2^*}(\omega_2) = \boldsymbol{D}_{011} = 0$ *in* $\boldsymbol{D}_{X_2=0,\overline{X}_3=1,C_2=c_2^*}(\omega_2) = \boldsymbol{D}_{000} = 0.$

Iz zgornjega izreka sledi, da je množica $\{X_2, \overline{X}_3\}$ *ireducibilna. Enako velja za množico* $\{\overline{X}_2, X_3\}$, *saj pri vrednosti* $c_2^* = 0$ *od* $C_2 = X_1$ *veljata naslednja pogoja:*

(i) $\boldsymbol{D}_{\overline{X}_2=1,X_3=1,C_2=c_2^*}(\omega_2) = \boldsymbol{D}_{001} = 1,$

(ii) $\boldsymbol{D}_{\overline{X}_2=1,X_3=0,C_2=c_2^*}(\omega_2) = \boldsymbol{D}_{000} = 0$ *in* $\boldsymbol{D}_{\overline{X}_2=0,X_3=1,C_2=c_2^*}(\omega_2) = \boldsymbol{D}_{011} = 0.$

Iz zgornjega izreka sledi, da se ireducibilnost ohrani, če razširimo populacijo:

Posledica 4.3 *Če je* $\boldsymbol{B}$ *ireducibilna za* $\mathcal{D}(\boldsymbol{C}, \Omega)$, *potem je* $\boldsymbol{B}$ *ireducibilna za* $\mathcal{D}(\boldsymbol{C}, \Omega^*)$ *pri vsakem* $\Omega^* \supset \Omega$.

Dokaz: Če Ω izpolnjuje pogoje (i) in (ii) iz zgornjega izreka 4.2, potem jih izpolnjuje tudi Ω^*. $\square$

4.2 Ireducibilnost in minimalni zadostni vzroki

4.2.1 Ireducibilne množice B za $\mathcal{D}(\mathbf{C}, \Omega)$, kjer $|\mathbf{B}| = |\mathbf{C}|$

V naslednji trditvi bomo spoznali, da pojma ireducibilnosti in minimalnega zadostnega vzroka za nek ω^* sovpadata, če $|\mathbf{B}| = |\mathbf{C}|$.

Trditev 4.4 *Naj bo* $\boldsymbol{B} \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}))$ *in* $|\boldsymbol{B}| = |\boldsymbol{C}|$. *Množica* $\boldsymbol{B}$ *je ireducibilna za* $\mathcal{D}(\boldsymbol{C}, \Omega)$, *če in samo če je minimalen zadosten vzrok za nek* ω^* *glede na* $\boldsymbol{C}$.

Dokaz: Naj bo $|\mathbf{B}| = |\mathbf{C}|$. Če $\mathbf{C}_2 = \emptyset$, potem pogoj (i) v izreku 4.2 o karakterizaciji ireducibilnih množic velja, če in samo če je $\mathbf{B}$ zadosten vzrok za D pri ω^*; pogoj (ii) v izreku 4.2 pa velja, če in samo če je $\mathbf{B}$ minimalen zadosten vzrok za D pri ω^*. □

Iz zgornje trditve sledi, da je vsak minimalen zadosten vzrok velikosti $|\mathbf{C}|$ vsebovan v vsaki reprezentaciji zadostnega vzroka:

Posledica 4.5 *Če $\boldsymbol{B} \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}))$, $|\boldsymbol{B}| = |\boldsymbol{C}|$ in je $\boldsymbol{B}$ minimalen zadosten vzrok za D pri nekem $\omega^* \in \Omega$, potem $\boldsymbol{B} \in \mathfrak{B}$ za vsako reprezentacijo $(\boldsymbol{A}, \mathfrak{B})$ za $\mathcal{D}(\boldsymbol{C}, \Omega)$.*

Dokaz: Uporabimo zgornjo trditev ter dejstvo, da je $\mathbf{B} \in \dot{\mathbb{P}}(\mathbb{L}(\mathbf{C}))$ največje možne velikosti $|\mathbf{C}|$. □

4.2.2 Ireducibilne množice B za $\mathcal{D}(\mathbf{C}, \Omega)$, kjer $|\mathbf{B}| < |\mathbf{C}|$

Če je množica dogodkov $\mathbf{B}$ velikosti $|\mathbf{B}| < |\mathbf{C}|$, potem se pogoja ireducibilnosti in minimalnega zadostnega vzroka razlikujeta. Primerjajmo pogoja iz izreka 4.2 ter pogoja (3.1) in (3.3).

Točka (i) iz izreka 4.2 zahteva veljavnost enakosti $D_{\mathbf{B}=1,\ \mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 1$ za neko vrednost $\mathbf{c}_2^*$ in nek ω^*; definicija zadostnega vzroka za ω^* pa zahteva veljavnost tega pogoja pri vsaki vrednosti $\mathbf{c}_2^*$.

Točka (ii) iz izreka 4.2 zahteva obstoj ene same vrednosti $\mathbf{c}_2^*$ (in nekega ω^*), da za vse $L \in \mathbf{B}$ velja $D_{\mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0,\ \mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 0$; definicija minimalnega zadostnega vzroka (za ω^*) pa zahteva za vsako množico $L \in \mathbf{B}$ obstoj vrednosti $\mathbf{c}_2^L$, da velja pogoj $D_{\mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0,\ \mathbf{C}_2=\mathbf{c}_2^L}(\omega^*) = 0$.

Torej v splošnem ireducibilne množice niso minimalni zadostni vzroki in minimalni zadostni vzroki niso nujno ireducibilni.

Primer 4.2 *Poglejmo primer minimalnega zadostnega vzroka, ki ni ireducibilen ter primer ireducibilne množice, ki ni minimalen zadosten vzrok. Obravnavajmo primer 2.1, kjer $\boldsymbol{C} = \{X_1, X_2, X_3\}$ in je populacija sestavljena samo iz drugega osebka ω_2. Množica $\{X_1, X_2\}$ je minimalen zadosten vzrok glede na $\boldsymbol{C}$ za ω_2, saj*

$$D_{111}(2) = D_{110}(2) = 1 \text{ in } D_{011}(2) = D_{100}(2) = 0.$$

Ni pa ireducibilna za $\mathcal{D}(\boldsymbol{C}, \Omega)$, saj pogoj (ii) iz izreka 4.2 ni izpolnjen: za $X_3 = 0$ velja $D_{010}(2) = 1$ ter za $X_3 = 1$ velja $D_{101}(2) = 1$.

Množica $\{X_1\}$ pa je ireducibilna za ω_2, saj

$$D_{111}(2) = 1 \text{ in } D_{011}(2) = 0,$$

ampak ni zadosten pogoj, saj velja $D_{100}(2) = 0$.

Čeprav ireducibilna množica $\mathbf{B}$, za katero velja $|\mathbf{B}| < |\mathbf{C}|$, ni minimalen zadosten vzrok in niti zadosten vzrok, ima naslednjo lastnost. Vsebovana je v nekem zadostnem vzroku $\mathbf{B}^*$ velikosti $|\mathbf{C}|$; v vsaki raprezentaciji lahko najdemo podmnožico $\mathbf{B}_j$ zadostnega vzroka $\mathbf{B}^*$, ki vsebuje ireducibilno množico $\mathbf{B}$. To lastnost bomo spoznali v naslednji posledici.

Posledica 4.6 *Naj bo $\boldsymbol{C} = \boldsymbol{C}_1 \dot{\cup} \boldsymbol{C}_2$, $\boldsymbol{B} \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}_1))$. Če je $\boldsymbol{B}$ ireducibilna za $\mathcal{D}(\boldsymbol{C}, \Omega)$, potem obstaja množica $\boldsymbol{B}^* \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}))$, da $|\boldsymbol{B}^*| = |\boldsymbol{C}|$ ter v vsaki reprezentaciji $(\boldsymbol{A}, \mathfrak{B})$ za $\mathcal{D}(\boldsymbol{C}, \Omega)$ obstaja $\boldsymbol{B}_j \in \mathfrak{B}$, da $\boldsymbol{B} \subseteq \boldsymbol{B}_j \subseteq \boldsymbol{B}^*$.*

Iz ireducibilnosti $\mathbf{B}$ sledi torej obstoj množice $\mathbf{B}^*$ velikosti $|\mathbf{C}|$ ter obstoj vsaj ene množice $\mathbf{B}_j$ v vsaki reprezentaciji, da velja $\mathbf{B} \subseteq \mathbf{B}_j \subseteq \mathbf{B}^*$. V splošnem lahko najdemo več množic $\mathbf{B}_j$, ki ustrezajo temu pogoju.

Dokaz: Uporabimo izrek 4.2 in dobimo vrednosti $\mathbf{c}_2^*$ od $\mathbf{C}_2$ in $\omega^* \in \Omega$, da veljata pogoja (i) in (ii). Definirajmo $\mathbf{B}^* = \mathbf{B} \cup \mathbf{C}^{[\mathbf{c}_2^*]}$. Iz pogoja (i) sledi $D_{\mathbf{B}^*=1}(\omega^*) = 1$. Po definiciji reprezentacije zadostnega vzroka iz $D_{\mathbf{B}^*=1}(\omega^*) = 1$ sledi, da $(\bigwedge(\mathbf{B}_j))_{\mathbf{B}^*=1} = 1$ za nek j, kjer $\mathbf{B}_j \in \mathfrak{B}$. Torej $\mathbf{B}_j \subseteq \mathbf{B}^*$. □

4.3 Ohranjanje ireducibilnosti pri večjih množicah

V naslednjem poglavju nas zanima, ali se lastnost ireducibilnosti za $\mathcal{D}(\mathbf{C}, \Omega)$ ohranja, če spreminjamo množico $\mathbf{C}$ in populacijo Ω.

Če populacijo povečamo, se ireducibilnost jasno ohrani; če pa jo zmanjšamo se po posledici 4.3 ireducubilnost prav tako ohrani.

Če povečamo množico $\mathbf{C}$ v $\mathbf{C} \cup \mathbf{C}'$ se ireducibilnost v splošnem ne ohrani. Potrebujemo namreč še dodaten pogoj, ki nam to omogoča. Dodaten pogoj se nanaša na povezavo med spremenljivkami iz $\mathbf{C}'$ in $\mathbf{C}$: spremenljivke iz $\mathbf{C}$ ne smejo *kavzalno vplivati* na spremenljivke iz $\mathbf{C}'$. Poglejmo kaj to pomeni.

Definicija 4.7 *(Kavzalen vpliv) Množica spremenljivk $\boldsymbol{A}$ kavzalno vpliva na spremenljivko B, če je za vsak osebek ω vrednost $B_{\boldsymbol{A}=\boldsymbol{a}}(\omega)$ konstanta pri vsaki vrednosti $\boldsymbol{a}$.*

Množica spremenljivk $\mathbf{A}$ *kavzalno vpliva na množico spremenljivk* $\mathbf{B}$, če $\mathbf{A}$ kavzalno vpliva na vsak element B iz $\mathbf{B}$.

Predpostavimo, da velja naslednji sklep: če množica $\mathbf{C}$ ne vpliva kavzalno na množico $\mathbf{C}'$, potem velja *aksiom relativizirane doslednosti* (relativized consistency axiom):

$$D_{\mathbf{C}=\mathbf{c},\mathbf{C}'=\mathbf{C}'(\omega)}(\omega) = D_{\mathbf{C}=\mathbf{c}}(\omega).$$

Poglejmo, kaj sledi iz izreka 4.2:

Trditev 4.8 *Naj bo $\boldsymbol{C} = \boldsymbol{C}_1 \dot\cup \boldsymbol{C}_2$, $\boldsymbol{B} \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}_1))$, $|\boldsymbol{B}| = |\boldsymbol{C}_1|$. Če je množica dogodkov $\boldsymbol{B}$ ireducibilna za $\mathcal{D}(\boldsymbol{C}, \Omega)$, $\boldsymbol{C}' \cap \boldsymbol{C} = \emptyset$ in $\boldsymbol{C}$ ne vpliva kavzalno na $\boldsymbol{C}'$, potem je $\boldsymbol{B}$ ireducibilna za $\mathcal{D}(\boldsymbol{C} \cup \boldsymbol{C}', \Omega)$.*

Dokaz: Iz izreka 4.2 vemo, da obstaja $\omega^* \in \Omega$ in vrednost $\mathbf{c}_2^*$ od $\mathbf{C}_2$, da veljata pogoja (i) in (ii). Naj bo $\mathbf{c}' = \mathbf{C}'(\omega)$. Spremenljivke $\mathbf{C}$ ne vplivajo kavzalno na $\mathbf{C}'$, torej pri poljubni vrednosti $\mathbf{b}$ velja

$$D_{\mathbf{B}=\mathbf{b},\mathbf{C}_2=\mathbf{c}_2^*,\mathbf{C}'=\mathbf{c}'}(\omega)^* = D_{\mathbf{B}=\mathbf{b},\mathbf{C}_2=\mathbf{c}_2^*,\mathbf{C}'=\mathbf{C}'(\omega^*)}(\omega)^* = D_{\mathbf{B}=\mathbf{b},\mathbf{C}_2=\mathbf{c}_2^*}(\omega)^*.$$

Posledično vrednosti ω^*, $\mathbf{c}_2^*$ in $\mathbf{c}'$ zadoščajo pogojema (i) in (ii) iz izreka 4.2 glede na $\mathbf{C} \cup \mathbf{C}'$. □

V zgornji posledici je predpostavka o kavzalnem vplivu nujno potrebna: če $\mathbf{C}$ kavzalno vpliva na $\mathbf{C}'$, lahko pri nekem osebku ω velja $\mathbf{C}'(\omega) \neq (\mathbf{C}')_{\mathbf{C}=\mathbf{c}}(\omega)$ pri določeni vrednosti $\mathbf{c}$. Posledično so tudi vrednosti možnih izidov $D_{\mathbf{C}=\mathbf{c},\mathbf{C}'=\mathbf{C}'(\omega)}(\omega)$ in $D_{\mathbf{C}=\mathbf{c}}(\omega)$ različne. Poglejmo primer, kjer predpostavka o kavzalnem vplivu ne velja in se ireducibilnost ne ohrani pri večji množici.

Primer 4.3 *Naj bo $\boldsymbol{C} = \{X_1, X_2\}$, $\boldsymbol{C}' = \{X_3\}$ ter naj veljata naslednji enakosti:*

$$(X_3)_{X_1=x_1,X_2=x_2}(\omega) = x_1 \wedge x_2$$

$$D_{X_1=x_1,X_2=x_2,X_3=x_3}(\omega) = x_3,$$

za vse $\omega \in \Omega$. Iz prve enakosti sledi, da $\boldsymbol{C}$ kavzalno vpliva na $\boldsymbol{C'}$. Iz druge enakosti pa sledi, da ne velja aksiom relativizirane doslednosti. Množica $\boldsymbol{B} = \{X_1, X_2\}$ je ireducibilna za $\mathcal{D}(\boldsymbol{C}, \Omega)$, saj veljajo naslednje enakosti:

$$D_{X_1=1,X_2=1} = 1,\ D_{X_1=1,X_2=0} = 0,\ D_{X_1=0,X_2=1} = 0;$$

množica $\boldsymbol{B}$ ni pa ireducibilne za $\mathcal{D}(\boldsymbol{C} \cup \boldsymbol{C'}, \Omega)$, saj ne obstaja tak x_3^*, da bi za nek osebek ω veljale naslednje enakosti:

$$D_{X_1=1,X_2=1,X_3=x_3^*}(\omega) = 1,\ D_{X_1=1,X_2=0,X_3=x_3^*}(\omega) = 0,\ D_{X_1=0,X_2=1,X_3=x_3^*}(\omega) = 0.$$

4.4 Zadosten pogoj ireducibilnosti pri pogoju monotonosti

V naslednjem razdelku bomo obravnavali zadosten pogoj ireducibilnosti pri pogoju monotonost, ki se poslužuje teorije grafov. Najprej bomo obravnavali nekaj osnovnih pojmov iz teorije grafov, potem pa bomo z njihovo pomočjo sestavili zadosten pogoj ireducibilnosti.

Poglejmo običajno definicijo grafa z razliko, da je množica vozlišč sestavljena iz dejanskih dogodkov.

Definicija 4.9 *(Graf). Graf $\mathfrak{B}$ na množici $\boldsymbol{B}$ je množica parov elementov iz $\boldsymbol{B}$,*

$$\mathfrak{B} \equiv \{\mathbf{E} \mid \mathbf{E} = \{B_1, B_2\} \subseteq \mathbf{B},\ B_1 \neq B_2\}.$$

Elemente množice $\mathfrak{B}$ imenujemo *povezave.*

Definicija 4.10 *(Povezana elementa). Elementa $L,\ L^* \in \boldsymbol{B}$ sta povezana v $\mathfrak{B}$, če obstaja zaporedje $L = L_1, \ldots, L_p = L^*$ različnih elementov v $\boldsymbol{B}$, da $\{L_i, L_{i+1}\} \in \mathfrak{B}$ za $i = 1, \ldots, p-1$.*

Zaporedju povezav med L in L^* pravimo *pot* v $\mathfrak{B}$.

Definicija 4.11 *(Drevo). Graf $\mathfrak{B}$ na $\boldsymbol{B}$ imenujemo drevo, če $|\mathfrak{B}| = |\boldsymbol{B}| - 1$ in je vsak par različnih elementov iz $\boldsymbol{B}$ povezan v $\mathfrak{B}$.*

Enostavna trditev iz teorije grafov zagotavlja, da obstaja v drevesu samo ena pot med poljubnima dvema elementoma.

Trditev 4.12 *Naj bo $\mathfrak{T}$ drevo na $\boldsymbol{B}$. Za vsak element $R \in \boldsymbol{B}$ obstaja naravna bijekcija med povezavami, kjer izključimo R, in drevesi:*

$$\phi_R^{\mathfrak{T}} : \boldsymbol{B} \setminus \{R\} \longrightarrow \mathfrak{T},$$

kjer $\phi_R^{\mathfrak{T}}(L) = \boldsymbol{E} = \{L, L'\}$ in $\boldsymbol{E} \in \mathfrak{B}$ je zadnja povezava na poti iz R v L.

Če za nek graf $\mathfrak{B}$ obstaja bijekcija iz zgornje trditve potem, je $\mathfrak{B}$ drevo, saj $|\mathfrak{B}| = |\mathbf{B}| - 1$ in za poljubna $L,\ L^* \in \mathbf{B}$ obstaja pot med njima, ki jo določa bijekcija.

V sledečih izrekih bomo spoznali, da s pomočjo drevesa na množici dogodkov, ki učinkujejo pozitivno, lahko sestavimo test ireducibilnosti. Zanimalo nas bo število različnih dreves na množici dogodokov, zato bomo uporabili naslednji izrek, ki ga ne bomo dokazali.

Izrek 4.13 (Cayley [2]). *Na množici $\boldsymbol{B}$ obstaja $|\boldsymbol{B}|^{|\boldsymbol{B}|-2}$ različnih dreves.*

Obravnavajmo izrek o zadostnem pogoju ireducibilnosti pri pogoju monotonosti, ki se poslužuje teorije grafov:

Izrek 4.14 *Naj bo $\boldsymbol{C} = \boldsymbol{C}_1 \dot{\cup} \boldsymbol{C}_2$, $\boldsymbol{B} = (\boldsymbol{B}_+ \dot{\cup} \boldsymbol{B}') \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}_1))$, $|\boldsymbol{B}| = |\boldsymbol{C}_1|$ in naj $\boldsymbol{B}_+$ pozitivno monotono učinkuje na D glede na $\boldsymbol{C}$. Če za neko drevo $\mathfrak{T}$ na $\boldsymbol{B}_+$, $\omega^* \in \Omega$ in neko vrednost $\boldsymbol{c}_2$, velja*

$$0 < D_{\boldsymbol{B}=\boldsymbol{1},\boldsymbol{C}_2=\boldsymbol{c}_2}(\omega^*) - \sum_{L \in \boldsymbol{B}} D_{\boldsymbol{B}\setminus\{L\}=\boldsymbol{1},L=0,\boldsymbol{C}_2=\boldsymbol{c}_2}(\omega^*) + \sum_{\boldsymbol{E}\in\mathfrak{T}} D_{\boldsymbol{B}\setminus\boldsymbol{E}=\boldsymbol{1},\boldsymbol{E}=\boldsymbol{0},\boldsymbol{C}_2=\boldsymbol{c}_2}(\omega^*), \tag{4.2}$$

potem je množica $\boldsymbol{B}$ ireducibilna za $\mathcal{D}(\boldsymbol{C},\Omega)$.

Če X negativno monotono učinkuje na D, potem lahko s pomočjo zgornjega izreka sestavimo teste ireducibilnosti množic, ki vsebujejo $\overline{X}$.

Dokaz: Zadostuje pokazati, da pri predpostavki monotonosti $\mathbf{B}_+$ iz zgornje neenakosti (4.2) sledi pogoj, ki je enakovreden definiciji ireducibilnosti tj. pogoj (4.1). Predpostavimo, da pogoj (4.1) ne velja, torej za vse vrednosti $\mathbf{c}_2$ in $\omega \in \Omega$ velja naslednja neenakost:

$$D_{\mathbf{B}=\mathbf{1},\ \mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) - \sum_{L\in\mathbf{B}} D_{\mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0,\ \mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) \leq 0.$$

Za vsak $\omega^* \in \Omega$ obstaja $R \in \mathbf{B}_+$, da velja

$$D_{\mathbf{B}=\mathbf{1},\ \mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) - \sum_{L\in\mathbf{B}'\cup\{R\}} D_{\mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0,\ \mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) \leq 0.$$

Namreč če pri osebku ω^* velja $D_{\mathbf{B}=\mathbf{1},\ \mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 0$, potem je zgornja neenačba veljavna. Če pa $D_{\mathbf{B}=\mathbf{1},\ \mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 1$, potem razlikujemo dva primera: če za $L \in \mathbf{B}'$ velja $D_{\mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0,\ \mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 1$, potem je neenačba ravno tako veljavna, če pa $D_{\mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0,\ \mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 0$ za vse $L \in \mathbf{B}'$, potem obstaja $R \in \mathbf{B}^+$, da velja $D_{\mathbf{B}\setminus\{R\}=\mathbf{1},\ R=0,\ \mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 1$.

Naj bo $\mathfrak{T}$ drevo na $\mathbf{B}_+$. Izpišimo desno stran enačbe (4.2) iz izjave zgornjega izreka, tako da prvo vsoto razstavimo v dve komponenti:

$$\begin{aligned}&\left[D_{\mathbf{B}=\mathbf{1},\mathbf{C}_2=\mathbf{c}_2}(\omega^*) - \sum_{L\in\mathbf{B}'\cup\{R\}} D_{\mathbf{B}\setminus\{L\}=\mathbf{1},L=0,\mathbf{C}_2=\mathbf{c}_2}(\omega^*)\right] \\ &+ \left[\sum_{L\in\mathbf{B}_+\setminus\{R\}} -D_{\mathbf{B}\setminus\{L\}=\mathbf{1},L=0,\mathbf{C}_2=\mathbf{c}_2}(\omega^*) + \sum_{\mathbf{E}\in\mathfrak{T}} D_{\mathbf{B}\setminus\mathbf{E}=\mathbf{1},\mathbf{E}=\mathbf{0},\mathbf{C}_2=\mathbf{c}_2}(\omega^*)\right].\end{aligned}$$

Ugotovili smo, da je prvi člen zgornjega izraza negativen; drugi člen pa zapišemo v naslednji obliki z uporabo trditve 5.12 o bijekciji med povezavami in drevesi:

$$\begin{aligned}&\sum_{L\in\mathbf{B}_+\setminus\{R\}} \{-D_{\mathbf{B}\setminus\{L\}=\mathbf{1},L=0,\mathbf{C}_2=\mathbf{c}_2}(\omega^*) + D_{\mathbf{B}\setminus\phi_{\mathbf{R}}^{\mathfrak{T}}(L)=\mathbf{1},\phi_{\mathbf{R}}^{\mathfrak{T}}(L)=\mathbf{0},\mathbf{C}_2=\mathbf{c}_2}(\omega^*)\} \\ &= \sum_{L\in\mathbf{B}_+\setminus\{R\}} \{-D_{\mathbf{B}\setminus\{L,L'\}=\mathbf{1},L'=1,L=0,\mathbf{C}_2=\mathbf{c}_2}(\omega^*) + D_{\mathbf{B}\setminus\{L,L'\}=\mathbf{1},\{L,L'\}=\mathbf{0},\mathbf{C}_2=\mathbf{c}_2}(\omega^*)\},\end{aligned}$$

kjer je $\phi_R^{\mathfrak{T}}(L) = \mathbf{E} = \{L, L'\}$ zadnja povezava na poti iz R v L. Dogodek L' ima pozitiven monoton učinek na D glede na $\mathbf{C}$, torej so vsi členi v zgornji vsoti negativni. Sledi, da pogoj (4.2) ne velja za poljubno drevo $\mathfrak{T}$.

$\square$

Primer 4.4 *Če $\boldsymbol{B} = \{X_1, X_2\} = \boldsymbol{B}_+ = \boldsymbol{C}$, potem je edino drevo na $\boldsymbol{B}_+$ sestavljeno iz povezave $\{X_1, X_2\}$. Če X_1 in X_2 pozitivno monotono učinkujeta na D (glede na $\boldsymbol{C}$) in velja naslednja neenakost za nek $\omega \in \Omega$,*

$$D_{11}(\omega) - (D_{10}(\omega) + D_{01}(\omega)) + D_{00}(\omega) > 0,$$

potem iz izreka 5.14 sledi, da je $\{X_1, X_2\}$ ireducibilna za $\mathcal{D}(\boldsymbol{C}, \Omega)$.

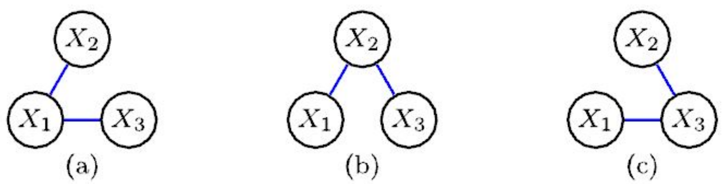
Primer 4.5 *Če $\boldsymbol{B} = \{X_1, X_2, X_3\} = \boldsymbol{B}_+ = \boldsymbol{C}$, potem sestavimo tri drevesa na $\boldsymbol{B}_+$, ki si jih lahko ogledamo na sliki 1. Pogoji, ki jih dobimo iz neenakosti (4.2), tako da upoštevamo drevesa na $\boldsymbol{B}_+$, so naslednji:*

$$\text{(a) } D_{111}(\omega) - (D_{110}(\omega) + D_{101}(\omega) + D_{101}(\omega)) + (D_{010}(\omega) + D_{001}(\omega)) > 0,$$

$$\text{(b) } D_{111}(\omega) - (D_{110}(\omega) + D_{101}(\omega) + D_{101}(\omega)) + (D_{100}(\omega) + D_{001}(\omega)) > 0,$$

$$\text{(c) } D_{111}(\omega) - (D_{110}(\omega) + D_{101}(\omega) + D_{101}(\omega)) + (D_{100}(\omega) + D_{010}(\omega)) > 0.$$

Množica $\boldsymbol{B}$ je ireducibilna za $\mathcal{D}(\boldsymbol{C}, \Omega)$, če pri nekem $\omega \in \Omega$ velja vsaj eden izmed zgornjih pogojev.



Slika 1: Drevesa na $\{X_1, X_2, X_3\}$

5 Testi

V naslednjem razdelku bomo obravnavali pogoje, ki nam omogočajo testiranje ireducibilnosti in minimalnosti v empiričnih primerih. Uporabljali bomo pogoj ireducibilnosti (4.1) ter pogoj iz izreka 4.14.

Spomnimo se pogoja ireducibilnosti (4.1): $\mathbf{B}$ je ireducibilna za $\mathcal{D}(\mathbf{C}, \Omega)$, če in samo če obstajajo $\omega^* \in \Omega$ in vrednosti $\mathbf{c}_2^*$ za $\mathbf{C}_2$, da velja naslednja neenakost:

$$D_{\mathbf{B}=\mathbf{1},\ \mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) - \sum_{L \in \mathbf{B}} D_{\mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0,\ \mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) > 0.$$

Zaradi osnovnega problema kavzalnega sklepanja ne moremo preveriti veljavnosti pogoja v konkretnih primerih. Namreč, nimamo na razpolago podatkov o vseh možnih izidih pri določenem osebku, ampak lahko opazujemo le en možen izid.

Poskusimo sestaviti test ireducibilnosti s pomočjo matematičnega upanja: nadomestimo možne izide $D_{\mathbf{C}=\mathbf{c}}(\omega)$ z matematičnim upanjem spremenljivke $D_{\mathbf{C}=\mathbf{c}}$, oziroma z

$$E[D_{\mathbf{C}=\mathbf{c}}] = \sum_{\omega \in \Omega} dP[D_{\mathbf{C}=\mathbf{c}}(\omega) = d],$$

kjer $P[B]$ označuje verjetnost dogodka B. Tudi pogoja

$$E[D_{\mathbf{B}=\mathbf{1},\ \mathbf{C}_2=\mathbf{c}_2^*}] - \sum_{L\in\mathbf{B}} E[D_{\mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0,\ \mathbf{C}_2=\mathbf{c}_2^*}] > 0$$

ne moremo preveriti, saj nimamo zadostnih podatkov o možnih izidih. Radi bi izračunali matematično upanje možnih izidov $E[D_{\mathbf{C}=\mathbf{c}}]$ s pomočjo pogojnega matematičnega upanja $E[D \mid \mathbf{C} = \mathbf{c}]$. Ti dve matematični upanji nista nujno enaki. Poglejmo primer, ko se razlikujeta.

Primer 5.1 *Vrnimo se k osnovnemu modelu možnih izidov iz poglavja 2.1, kjer X označuje izpostavljenost zdravljenju, Ω_1 zdravljeno skupino, Ω_0 kontrolno skupino in spremenljivka D zdravstveno stanje. Če je podpopulacija Ω_1 sestavljena iz osebkov šibkega zdravja, podpopulacija Ω_0 pa iz osebkov krepkega zdravja, bo povprečno zdravstveno stanje osebkov v Ω_1,*

$$E[D \mid X = 1] = \sum_{\omega\in\Omega_1} dP[D(\omega) = d],$$

slabše od povprečnega zdravstvenega stanja osebkov celotne populacije Ω, ko bi bili ti izpostavljeni zdravljenju,

$$E[D_{X=1}] = \sum_{\omega\in\Omega} dP[D_{X=x}(\omega) = d].$$

Spoznali bomo, da matematični upanji $E[D_{\mathbf{C}=\mathbf{c}}]$ in $E[D \mid \mathbf{C} = \mathbf{c}]$ sovpadata v idealnem primeru, ko sta spremenljivki $D_{\mathbf{C}=\mathbf{c}}$ in $\mathbf{C}$ neodvisni. To se v empiričnih primerih poredkoma pripeti. Pogosto moramo poiskati manjše podpopulacije, znotraj katerih sta spremenljivki neodvisni. Podpopulacije karakteriziramo s pomočjo *pojasnjevalnih spremenljivk*, ki označujejo lastnosti podpopulacije. Zavzemajo vrednosti iz poljubnega prostora R in so definirane na celotni populaciji. Označevali bomo $\mathbf{W} = (W_1, \ldots, W_m)$ vektor pojasnjevalnih spremenljivk. Vsaka vrednost $\mathbf{w} = (w_1, \ldots, w_m) \in R^m$ karakterizira podpopulacijo

$$\Omega_{\mathbf{w}} = \{\omega \in \Omega \mid \mathbf{W}(\omega) = \mathbf{w}\}.$$

Če sta spremenljivki $D_{\mathbf{C}=\mathbf{c}}$ in $\mathbf{C}$ odvisni, pravimo, da ima $\mathbf{C}$ *zavajajoč učinek* (confounding effect) na D. Spoznali bomo, da pri predpostavki neodvisnosti spremenljivk $D_{\mathbf{C}=\mathbf{c}}$ in $\mathbf{C}$ v določeni podpopulaciji $\Omega_{\mathbf{w}}$ lahko oblikujemo naslednji test ireducibilnosti za osebke iz $\Omega_{\mathbf{w}}$:

$$E[D \mid \mathbf{B} = \mathbf{1},\ \mathbf{C}_2 = \mathbf{c}_2^*] - \sum_{L\in\mathbf{B}} E[D \mid \mathbf{B} \setminus \{L\} = \mathbf{1},\ L = 0,\ \mathbf{C}_2 = \mathbf{c}_2^*] > 0$$

5.1 Popravljanje zavajajočega učinka.

Povedali smo, da sta običajno spremenljivki $D_{\mathbf{C}=\mathbf{c}}$ in $\mathbf{C}$ neodvisni le v manjših podpopulacijah, ki jih označuje vektor spremenljivk $\mathbf{W}$. V tem primeru pravimo, da množica pojasnjevalnih spremenljivk $\mathbf{W}$ *popravlja zavajajoč učinek* $\mathbf{C}$ na D. Poglejmo formalno definicijo. Označimo z $A \perp\!\!\!\perp B \mid C$ neodvisnost slučajnih spremenljivk A in B pogojno na spremenljivko C.

Definicija 5.1 *(Popravljanje zavajajočega učinka). Množica pojasnjevalnih spremenljivk $\boldsymbol{W}$ popravlja zavajajoč učinek $\boldsymbol{C}$ na D, če*

$$D_{\boldsymbol{C}=\boldsymbol{c}} \perp\!\!\!\perp \boldsymbol{C} \mid \boldsymbol{W} = \boldsymbol{w} \tag{5.1}$$

za vse $\boldsymbol{c}$, $\boldsymbol{w}$.

Pravimo tudi, da pri vrednosti $\mathbf{w}$ od $\mathbf{W}$ množica $\mathbf{C}$ ne *učinkuje zavajajoče* na D. Če torej množica $\mathbf{W}$ popravlja zavajajoč učinek $\mathbf{C}$ na D, sta spremenljivki $D_{\mathbf{C}=\mathbf{c}}$ in $\mathbf{C}$ neodvisni znotraj podpopulacije $\Omega_{\mathbf{w}}$.

Neodvisnost spremenljivk znotraj podpopulacije $\Omega_{\mathbf{w}}$ lahko dosežemo z naključnim določanjem vzorcev znotraj podpopulacije. Naključno določanje vzorcev igra zelo pomembno vlogo pri statističnem sklepanju. Če sklepamo iz podatkov, ki smo jih nepremišljeno izbrali, lahko pridemo do napačnih rezultatov, ki ne izražajo resnične kavzalne zveze.

Poglejmo, kako odsotnost zavajajočega učinka vpliva na pogojno matematično upanje možnih izidov.

Trditev 5.2 *Naj $\boldsymbol{W}$ popravlja zavajajoč učinek $\boldsymbol{C}$ na D in naj bo $P(\boldsymbol{C}=\boldsymbol{c}, \boldsymbol{W}=\boldsymbol{w})>0$, potem*

$$E[D_{\boldsymbol{C}=\boldsymbol{c}} \mid \boldsymbol{W}=\boldsymbol{w}] = E[D \mid \boldsymbol{C}=\boldsymbol{c},\ \boldsymbol{W}=\boldsymbol{w}].$$

Dokaz: Najprej posplošimo enačbo (2.1) in izrazimo izid D s pomočjo možnih izidov $D_{\mathbf{C}=\mathbf{c}'}$:

$$D = \sum_{\mathbf{c}'\in\{0,1\}^s} \bigwedge(\mathbf{B}^{[\mathbf{c}']}) D_{\mathbf{C}=\mathbf{c}'},$$

kjer je $\mathbf{B}^{[\mathbf{c}']}$ množica dejanskih dogodkov pri vrednosti $\mathbf{c}'$ od $\mathbf{C}$. Veljajo naslednje enakosti:

$$\begin{aligned}
E[D \mid \mathbf{C}=\mathbf{c}, \mathbf{W}=\mathbf{w}] &= \sum_{\mathbf{c}'\in\{0,1\}^s} E\big[\bigwedge(\mathbf{B}^{[\mathbf{c}']}) D_{\mathbf{C}=\mathbf{c}'} \mid \mathbf{C}=\mathbf{c},\ \mathbf{W}=\mathbf{w}\big] \\
&= \sum_{\mathbf{c}'\in\{0,1\}^s} E\big[\bigwedge(\mathbf{B}^{[\mathbf{c}']}) \mid \mathbf{C}=\mathbf{c},\ \mathbf{W}=\mathbf{w}] E[D_{\mathbf{C}=\mathbf{c}'} \mid \mathbf{C}=\mathbf{c},\ \mathbf{W}=\mathbf{w}\big] \\
&= E[D_{\mathbf{C}=\mathbf{c}} \mid \mathbf{W}=\mathbf{w}],
\end{aligned}$$

kjer pri drugi enakosti uporabimo pogojno neodvisnost $D_{\mathbf{C}=\mathbf{c}}$ in $\mathbf{C}=\mathbf{c}$, iz katere sledi pogojna neodvisnost $D_{\mathbf{C}=\mathbf{c}}$ in $\mathbf{B}^{[\mathbf{c}]}$. Pri tretji enakosti pa uporabimo dejstvo, da $(\bigwedge(\mathbf{B}^{[\mathbf{c}']}))_{\mathbf{C}=\mathbf{c}}=1$, če in samo če $\mathbf{c}'=\mathbf{c}$. □

Opomba 5.1 *Če $\boldsymbol{W}$ popravlja zavajajoč učinek $\boldsymbol{C}$ na D, potem $\boldsymbol{W}$ zadošča tudi pri popravljanju zavajajočega učinka $\boldsymbol{B}$ na D, kjer $\boldsymbol{B}\in\dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}))$, $|\boldsymbol{B}|=|\boldsymbol{C}|$.*

5.1.1 Zavajajoč učinek pri računanju povprečnega kavzalnega učinka

V naslednjem poglavju bomo poudarili pomembnost popravljanja zavajajočega učinka v statističnem sklepanju. Predstavili bomo zavajajoča učinka, ki se pojavita pri računanju kavzalnega učinka.

Primer 5.2 *Obravnavajmo spet osnovni model možnih izidov iz poglavja 2.1, kjer X označuje izpostavljenost zdravljenju, Ω_1 zdravljeno skupino, Ω_0 kontrolno skupino in spremenljivka D zdravstveno stanje. Če $\boldsymbol{W}$ popravlja zavajajoč učinek X na D, potem je povprečni učinek zdravljenja v podskupini $\Omega_{\boldsymbol{W}=\boldsymbol{w}}$ enak*

$$\begin{aligned}
E[\delta \mid \boldsymbol{W}=\boldsymbol{w}] &= E[D_1 - D_0 \mid \boldsymbol{W}=\boldsymbol{w}] \\
&= E[D_1 \mid \boldsymbol{W}=\boldsymbol{w}] - E[D_0 \mid \boldsymbol{W}=\boldsymbol{w}] \\
&= E[D \mid X=1, \boldsymbol{W}=\boldsymbol{w}] - E[D \mid X=0, \boldsymbol{W}=\boldsymbol{w}] \\
&= E[D_1 \mid X=1, \boldsymbol{W}=\boldsymbol{w}] - E[D_0 \mid X=0, \boldsymbol{W}=\boldsymbol{w}],
\end{aligned}$$

kjer predzadnja enakost sledi iz trditve 5.2. V tem primeru lahko povprečni kavzalni učinek izračunamo s pomočjo opazovanih podatkov. Člen $E[D_1 \mid X=1, \boldsymbol{W}=\boldsymbol{w}] - E[D_0 \mid X=0, \boldsymbol{W}=\boldsymbol{w}]$ imenujemo naiven povprečni učinek zdravljenja v podpopulaciji $\Omega_{\boldsymbol{W}=\boldsymbol{w}}$.

Če $\boldsymbol{W}$ ne popravlja zavajajočega učinka X na D, zgornja enačba ne velja in moramo upoštevati še dodatne zavajajoče člene. Razstavimo najprej povprečni kavzalni učinek na sledeče komponente:

$$\begin{aligned}
E[\delta \mid \boldsymbol{W}=\boldsymbol{w}] =& \{\sum_{\omega\in\Omega_1} D_1(\omega)P[D_1 = D_1(\omega)] + \sum_{\omega\in\Omega_0} D_1(\omega)P[D_1 = D_1(\omega)]\} \\
&- \{\sum_{\omega\in\Omega_1} D_0(\omega)P[D_0 = D_0(\omega)] + \sum_{\omega\in\Omega_0} D_0(\omega)P[D_0 = D_0(\omega)]\} \\
=& \{\pi E[D_1 \mid X=1, \boldsymbol{W}=\boldsymbol{w}] + (1-\pi)E[D_1 \mid X=0, \boldsymbol{W}=\boldsymbol{w}]\} \\
&- \{\pi E[D_0 \mid X=1, \boldsymbol{W}=\boldsymbol{w}] + (1-\pi)E[D_0 \mid X=0, \boldsymbol{W}=\boldsymbol{w}]\},
\end{aligned}$$

kjer je π delež populacije v zdravljeni skupini. Če desni strani enačbe prištejemo in odštejemo izraz $E[D_1 \mid X=1, \boldsymbol{W}=\boldsymbol{w}]$, sledi:

$$\begin{aligned}
E[\delta \mid \boldsymbol{W}=\boldsymbol{w}] =& E[D_1 \mid X=1, \boldsymbol{W}=\boldsymbol{w}] - E[D_0 \mid X=0, \boldsymbol{W}=\boldsymbol{w}] \\
&- (1-\pi)E[D_1 \mid X=1, \boldsymbol{W}=\boldsymbol{w}] + \pi E[D_0 \mid X=0, \boldsymbol{W}=\boldsymbol{w}] \\
&+ (1-\pi)E[D_0 \mid X=0, \boldsymbol{W}=\boldsymbol{w}].
\end{aligned}$$

Če prištejemo in odštejemo še izraz $E[D_0 \mid X=1, \boldsymbol{W}=\boldsymbol{w}] + E[D_0 \mid X=0, \boldsymbol{W}=\boldsymbol{w}]$, dobimo naslednjo enačbo:

$$\begin{aligned}
E[\delta \mid \boldsymbol{W}=\boldsymbol{w}] =& E[D_1 \mid X=1, \boldsymbol{W}=\boldsymbol{w}] - E[D_0 \mid X=0, \boldsymbol{W}=\boldsymbol{w}] \\
&- \{E[D_0 \mid X=1, \boldsymbol{W}=\boldsymbol{w}] - E[D_0 \mid X=0, \boldsymbol{W}=\boldsymbol{w}]\} \\
&- (1-\pi)\{E[\delta \mid X=1, \boldsymbol{W}=\boldsymbol{w}] - E[\delta \mid X=0, \boldsymbol{W}=\boldsymbol{w}]\}.
\end{aligned}$$

Zadnja dva člena iz desne strani zgornje enačbe predstavljata zavajajoč učinek pri ocenjevanju povprečnega kavzalnega učinka: člen $E[D_0 \mid X=1, \boldsymbol{W}=\boldsymbol{w}] - E[D_0 \mid X=0, \boldsymbol{W}=\boldsymbol{w}]$ predstavlja t. i. učinek izbora (selection bias); člen $E[\delta \mid X=1, \boldsymbol{W}=\boldsymbol{w}] - E[\delta \mid X=0, \boldsymbol{W}=\boldsymbol{w}]$ pa predstavlja pristranskost, ki izhaja iz razlike učinka zdravljenja.

Poglejmo predpostavki, ki zagotavljata, da je povprečen učinek zdravljenja enak naivnemu povprečnemu učinku zdravljenja:

$$\begin{aligned}
Predpostavka\ 1: \ E[D_1 \mid X=1, \boldsymbol{W}=\boldsymbol{w}] &= E[D_1 \mid X=0, \boldsymbol{W}=\boldsymbol{w}] \\
Predpostavka\ 2: \ E[D_0 \mid X=1, \boldsymbol{W}=\boldsymbol{w}] &= E[D_0 \mid X=0, \boldsymbol{W}=\boldsymbol{w}]
\end{aligned}$$

Prva predpostavka obravnava možne izide v podpopulaciji, druga pa postopek prirejanja zdravljenja. V primeru veljavnosti obeh predpostavk je naiven povprečni učinek zdravljenja v podpopulaciji enak bodisi ATC-ju kot ATT-ju v podpopulaciji, ki smo ju definirali v poglavju 2.1. Predpostavki sta veljavni pri naključnem določanju vzorca, saj sta v tem primeru X in (D_0, D_1) pogojno neodvisna.

V primeru, da je prva predpostavka veljavna ter druga neveljavna, je naiven povprečni učinek zdravljenja v podpopulaciji enak ATC v podpopulaciji, saj velja $E[D_1 \mid X=1] - E[D_0 \mid X=0] = E[D_1 \mid X=0] - E[D_0 \mid X=0]$. V primeru pa, da je druga predpostavka veljavna ter prva neveljavna, je naiven povprečni učinek zdravljenja v podpopulaciji enak ATT v podpopulaciji, saj velja $E[D_1 \mid X=1] - E[D_0 \mid X=0] = E[D_1 \mid X=1] - E[D_0 \mid X=1]$.

5.2 Testi ireducibilnosti

Formulirajmo najprej teste ireducibilnosti brez pogoja monotonosti.

Izrek 5.3 *Naj bo $\boldsymbol{C} = \boldsymbol{C}_1 \dot{\cup} C_2$, $\boldsymbol{B} \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}_1))$, $|\boldsymbol{B}| = |\boldsymbol{C}_1|$. Če $\boldsymbol{W}$ popravlja zavajajoč učinek $\boldsymbol{C}$ na D in za nek $\boldsymbol{c}_2, \boldsymbol{w}$ velja*

$$\begin{aligned} 0 <& E[D \mid \boldsymbol{B} = \boldsymbol{1}, \boldsymbol{C}_2 = \boldsymbol{c}_2, \boldsymbol{W} = \boldsymbol{w}] \\ &- \sum_{L \in \boldsymbol{B}} E[D \mid \boldsymbol{B} \setminus \{L\} = \boldsymbol{1},\ L = 0,\ \boldsymbol{C}_2 = \boldsymbol{c}_2,\ \boldsymbol{W} = \boldsymbol{w}], \end{aligned} \tag{5.2}$$

potem je $\boldsymbol{B}$ ireducibilna množica za $\mathcal{D}(\boldsymbol{C}, \Omega)$.

Dokaz: Predpostavimo, da $\mathbf{B}$ ni ireducibilna za $\mathcal{D}(\mathbf{C}, \Omega)$. Iz izreka 4.2 sledi naslednja neenakost za vse $\omega \in \Omega$ in $\mathbf{c}_2$:

$$D_{\mathbf{B}=\mathbf{1}, \mathbf{C}_2=\mathbf{c}_2}(\omega) - \sum_{L \in \mathbf{B}} D_{\mathbf{B} \setminus \{L\}=\mathbf{1}, L=0, \mathbf{C}_2=\mathbf{c}_2}(\omega) \leq 0.$$

Torej za vsak $\mathbf{w}$ velja:

$$E\left[D_{\mathbf{B}=\mathbf{1}, \mathbf{C}_2=\mathbf{c}_2} - \sum_{L \in \mathbf{B}} D_{\mathbf{B} \setminus \{L\}=\mathbf{1}, L=0, \mathbf{C}_2=\mathbf{c}_2} \middle| \mathbf{W} = \mathbf{w} \right] \leq 0.$$

Če uporabimo trditev 5.2 dobimo negacijo pogoja (5.2):

$$E[D|\mathbf{B} = \mathbf{1}, \mathbf{C}_2 = \mathbf{c}_2, \mathbf{W} = \mathbf{w}] - \sum_{L \in \mathbf{B}} E[D|\mathbf{B} \setminus \{L\} = \mathbf{1}, L = 0, \mathbf{C}_2 = \mathbf{c}_2, \mathbf{W} = \mathbf{w}] \leq 0.$$

$\square$

Opomba 5.2 *Iz trditve 4.8 sledi, da pri pogoju (5.2) je $\boldsymbol{B}$ ireducibilen tudi za $\mathcal{D}(\boldsymbol{C} \cup \boldsymbol{C}', \Omega)$, če spremenljivke $\boldsymbol{C}$ ne kavzalno vplivajo na spremenljivke iz $\boldsymbol{C}'$.*

5.3 Testi ireducibilnosti pri pogoju monotonosti

Poglejmo, kako monotonost vpliva na teste ireducibilnosti. Obravnavali bomo test, ki sledi iz izreka 4.14. Spoznali bomo, da je bolj učinkovit od pogoja iz izreka 5.3 ter ni toliko zahteven.

Izrek 5.4 *Naj bo $\boldsymbol{C} = \boldsymbol{C}_1 \dot{\cup} \boldsymbol{C}_2$, $\boldsymbol{B} = (\boldsymbol{B}_+ \cup \boldsymbol{B}') \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}_1))$, $|\boldsymbol{B}| = |\boldsymbol{C}_1|$. Naj $\boldsymbol{B}_+$ pozitivno monotono učinkuje na D glede na $\boldsymbol{C}$. Naj $\boldsymbol{W}$ zadošča pri popravljanju zavajajočega učinka $\boldsymbol{C}$ na D. Če za neko drevo $\mathfrak{T}$ na $\boldsymbol{B}_+$ in vrednosti $\boldsymbol{c}_2$, $\boldsymbol{w}$ velja*

$$\begin{aligned} 0 <& E[D \mid \boldsymbol{B} = \boldsymbol{1}, \boldsymbol{C}_2 = \boldsymbol{c}_2, \boldsymbol{W} = \boldsymbol{w}] \\ &- \sum_{L \in \boldsymbol{B}} E[D \mid \boldsymbol{B} \setminus \{L\} = \boldsymbol{1}, L = 0, \boldsymbol{C}_2 = \boldsymbol{c}_2, \boldsymbol{W} = \boldsymbol{w}] \\ &+ \sum_{\boldsymbol{E} \in \mathfrak{T}} E[D \mid \boldsymbol{B} \setminus \boldsymbol{E} = \boldsymbol{1}, \boldsymbol{E} = \boldsymbol{0}, \boldsymbol{C}_2 = \boldsymbol{c}_2, \boldsymbol{W} = \boldsymbol{w}], \end{aligned} \tag{5.3}$$

potem je množica $\boldsymbol{B}$ ireducibilna za $\mathcal{D}(\boldsymbol{C}, \Omega)$.

Dokaz: Podoben dokazu izreka 5.3. $\square$

Če vsak element iz $\mathbf{B}$ pozitivno monotono učinkuje na D, potem iz izreka 4.13 sledi obstoj $|\mathbf{B}|^{|\mathbf{B}|-2}$ zadostnih pogojev ireducibilnosti množice $\mathbf{B}$.

Opomba 5.3 *Dodajmo še, da iz trditve 4.8 sledi ireducibilnost $\boldsymbol{B}$ tudi za $\mathcal{D}(\boldsymbol{C} \cup \boldsymbol{C}', \Omega)$, če $\boldsymbol{C}$ ne vpliva kavzalno na $\boldsymbol{C}'$.*

5.4 Testi minimalnosti zadostnega vzroka

Spoznali smo, da se v primeru $|\mathbf{B}| = |\mathbf{C}|$ pojma ireducibilnosti ter minimalnega zadostnega vzroka skladata. Torej so testi ireducibilnosti iz prejšnjih podpoglavij tudi zadostni pogoji, da je $\mathbf{B}$ minimalen zadosten vzrok glede na $\mathbf{C}$.

Če pa velja $|\mathbf{B}| < |\mathbf{C}|$, potem testi iz prejšnjih poglavij niso zadostni pogoji minimalnosti. Sestavimo nove teste, ki so posebni primeri teh.

V naslednjem izreku bomo ugotovili, da pri predpostavki monotonosti pogoj iz izreka 5.3 pri $\mathbf{c}_2 = \mathbf{0}$ zadostuje pri preverjanju minimalnosti zadostnega vzroka.

Izrek 5.5 *Naj bo $\boldsymbol{C} = \boldsymbol{C}_1 \dot\cup \boldsymbol{C}_2$, $\boldsymbol{B} \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}_1))$, $|\boldsymbol{B}| = |\boldsymbol{C}_1|$. Naj $\boldsymbol{C}_2$ pozitivno monotono učinkuje na D glede na $\boldsymbol{C}$ in naj $\boldsymbol{W}$ zadošča pri popravljanju zavajajočega učinka $\boldsymbol{C}$ na D. Če za nek $\boldsymbol{w}$ velja pogoj*

$$\begin{aligned} 0 < & E[D \mid \boldsymbol{B} = \boldsymbol{1}, \boldsymbol{C}_2 = \boldsymbol{0}, \boldsymbol{W} = \boldsymbol{w}] \\ & - \sum_{L \in \boldsymbol{B}} E[D \mid \boldsymbol{B} \setminus \{L\} = \boldsymbol{1},\ L = 0,\ \boldsymbol{C}_2 = \boldsymbol{0},\ \boldsymbol{W} = \boldsymbol{w}], \end{aligned} \tag{5.4}$$

potem je $\boldsymbol{B}$ minimalen zadosten vzrok za D glede na $\boldsymbol{C}$ za nek osebek $\omega \in \Omega$.

Poudarimo, da je obravnavana množica $\mathbf{B}$ ireducibilna, saj je zgornji pogoj poseben primer izreka 5.3. Opazimo, da izrek zagotavlja minimalnost samo za enega osebka. Torej je pri predpostavki monotonega učinka ireducibilna množica tudi minimalen zadosten vzrok vsaj za nek osebek ω.

Zgornji pogoj je strožji od pogoja v izreku 5.3, saj zahteva veljavnost neenakosti pri točno določeni vrednosti $\mathbf{c}_2 = \mathbf{0}$.

Dokaz: Iz predpostavke sledi

$$E\left[D_{\mathbf{B}=\mathbf{1},\mathbf{C}_2=\mathbf{0},\mathbf{W}=\mathbf{w}} - \sum_{L \in \mathbf{B}} D_{\mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0,\ \mathbf{C}_2=\mathbf{0},\ \mathbf{W}=\mathbf{w}}\right] > 0,$$

torej obstaja vsaj en osebek ω^*, za katerega velja

$$D_{\mathbf{B}=\mathbf{1},\mathbf{C}_2=\mathbf{0}}(\omega^*) - \sum_{L \in \mathbf{B}} D_{\mathbf{B}\setminus\{L\}=\mathbf{1},L=0,\mathbf{C}_2=0}(\omega^*) > 0.$$

Torej $D_{\mathbf{B}=\mathbf{1},\mathbf{C}_2=\mathbf{0}}(\omega^*) = 1$ in $D_{\mathbf{B}\setminus\{L\}=\mathbf{1},L=0,\mathbf{C}_2=0}(\omega^*) = 0$ za vse $L \in \mathbf{B}$. Iz trditve 3.6 sledi, da je $\mathbf{B}$ minimalen zadosten vzrok za osebek ω^*. □

Izrek 5.6 *Naj bo $\boldsymbol{C} = \boldsymbol{C}_1 \dot\cup \boldsymbol{C}_2$, $\boldsymbol{B} = (\boldsymbol{B}_+ \dot\cup \boldsymbol{B}') \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}_1))$, $|\boldsymbol{B}| = |\boldsymbol{C}_1|$. Naj $\boldsymbol{B}_+ \cup \boldsymbol{C}_2$ pozitivno monotono učinkuje na D glede na $\boldsymbol{C}$. Naj $\boldsymbol{W}$ popravlja zavajajoč učinek $\boldsymbol{C}$ na D. Če za nek $\boldsymbol{w}$ in za neko drevo $\mathfrak{T}$ na $\boldsymbol{B}_+$ velja pogoj*

$$\begin{aligned} 0 < & E[D \mid \boldsymbol{B} = 1, \boldsymbol{C}_2 = \boldsymbol{0}, \boldsymbol{W} = \boldsymbol{w}] \\ & - \sum_{L \in \boldsymbol{B}} E[D \mid \boldsymbol{B} \setminus \{L\} = \boldsymbol{1}, L = 0, \boldsymbol{C}_2 = \boldsymbol{0}, \boldsymbol{W} = \boldsymbol{w}] \\ & + \sum_{\boldsymbol{E} \in \mathfrak{T}} E[D \mid \boldsymbol{B} \setminus \boldsymbol{E} = \boldsymbol{1}, \boldsymbol{E} = \boldsymbol{0}, \boldsymbol{C}_2 = \boldsymbol{0}, \boldsymbol{W} = \boldsymbol{w}], \end{aligned} \tag{5.5}$$

potem je $\boldsymbol{B}$ minimalen zadosten vzrok za D glede na $\boldsymbol{C}$ pri nekem $\omega \in \Omega$.

Dokaz: Podoben prejšnjemu dokazu. $\square$

6 Singularne interakcije in testi singularnosti

V naslednjem poglavju bomo obravnavali posebne interakcije med elementi minimalnega zadostnega vzroka. Za motivacijo poglejmo primer iz genetike.

V kontekstu dveh binarnih genetskih dejavnikov X_1 in X_2 pravimo, da je prisotna *kompozicijska epistaza* (compositional epistasis), če za nek ω^* velja

$$D_{11}(\omega^*) = 1 \text{ in } D_{10}(\omega^*) = D_{01}(\omega^*) = D_{00}(\omega^*) = 0.$$

Torej v primeru, da sta oba dejavnika X_1 in X_2 prisotna, se izid D pripeti, če pa je kateri od dejavnikov odsoten se njihov učinek prikrije in se izid D ne pripeti. Spomnimo se, da minimalen zadosten vzrok $\{X_1, X_2\}$ za D glede na $\{X_1, X_2\}$ pri ω^* izpolnjuje naslednje pogoje:

$$D_{11}(\omega^*) = 1 \text{ in } D_{10}(\omega^*) = D_{01}(\omega^*) = 0.$$

Pogoj $D_{00}(\omega^*) = 0$ pri minimalnosti ne velja nujno, torej je kompozicijska epistaza poseben primer minimalnega zadostnega vzroka. Primer kompozicijske epistaze bomo zaobjeli s pojmom *singularnega minimalnega zadostnega vzroka*: definirali bomo singularen vzrok kot minimalen zadosten vzrok, za katerega nujno velja tudi pogoj $D_{00}(\omega^*) = 0$.

6.1 Definicija ter potrebni in zadostni pogoj

Definicija 6.1 *Minimalen zadosten vzrok $\boldsymbol{B}$ za D glede na $\boldsymbol{C}$ pri ω^* je singularen, če ne obstaja nobena množica $\boldsymbol{B}' \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}))$, $\boldsymbol{B}' \neq \boldsymbol{B}$, ki je minimalen zadosten vzork za D glede na $\boldsymbol{C}$ pri ω^*. Množica $\boldsymbol{B}$ je singularna za $\mathcal{D}(\boldsymbol{C}, \Omega)$, če je $\boldsymbol{B}$ singularna glede na $\boldsymbol{C}$ za nek $\omega^* \in \Omega$.*

Če je $\mathbf{B}$ singularen vzrok, potem ne obstaja nobena množica $\mathbf{B}' \in \dot{\mathbb{P}}(\mathbb{L}(\mathbf{C}))$, $\mathbf{B}' \neq \mathbf{B}$, ki je zadosten vzrok za D glede na $\mathbf{C}$ pri ω^*, saj bi drugače vzeli njeno minimalno podmnožico. Če je množica $\mathbf{B}$ singularna za $\mathcal{D}(\mathbf{C}, \Omega)$, pravimo, da med njenimi elementi potekajo *singularne interakcije.*

V naslednjem izreku obravnavajmo zadosten in potreben pogoj singularnosti.

Izrek 6.2 *Naj bo $\boldsymbol{C} = \boldsymbol{C}_1 \dot\cup \boldsymbol{C}_2$, $\boldsymbol{B} \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}_1))$, $|\boldsymbol{B}| = |\boldsymbol{C}_1|$. Minimalen zadosten vzrok $\boldsymbol{B}$ je singularen za $\mathcal{D}(\boldsymbol{C}, \Omega)$, če in samo če obstaja $\omega^* \in \Omega$, da za vse vrednosti $\boldsymbol{c}_2^*$ in $\boldsymbol{b}$ velja naslednji pogoj:*

$$D_{\boldsymbol{B}=\boldsymbol{b}, \boldsymbol{C}_2=\boldsymbol{c}_2^*}(\omega^*) = 1 \iff \boldsymbol{b} = \boldsymbol{1}. \tag{6.1}$$

Singularnost zahteva, da je množica dogodkov $\mathbf{B}$ edini vzrok izida D: če se izid pripeti, so se torej nujno pripetili tudi dogodki iz $\mathbf{B}$. To lahko jasno vidimo iz zgornjega izreka.

Opomba 6.1 *Pogoj (6.1) velja za vse vrednosti $\boldsymbol{c}_2^*$ in $\boldsymbol{b}$, če in samo če obstaja osebek ω^*, da velja naslednji pogoj*

$$D_{\boldsymbol{C}=\boldsymbol{c}}(\omega^*) = \left(\bigwedge(\boldsymbol{B})\right)_{\boldsymbol{c}} \tag{6.2}$$

za vse $\boldsymbol{c}$.

Poglejmo dokaz izreka 6.2.

Dokaz: Naj bo $\mathbf{B}$ minimalen zadosten vzrok za D pri ω^*. Zadostnost vzroka $\mathbf{B}$ za D pri ω^* je ekvivalentna pogoju: $\mathbf{b} = \mathbf{1} \Rightarrow D_{\mathbf{B}=\mathbf{b},\mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 1$. Treba je torej še pokazati, da je singularnost minimalnega zadostnega vzroka ekvivalentna pogoju: $D_{\mathbf{B}=\mathbf{b},\mathbf{C}_2=\mathbf{c}_2^*}(\omega)^* = 1 \Rightarrow \mathbf{b} = \mathbf{1}$. Pokažimo, da ne obstaja noben minimalen zadosten vzrok od D pri ω^*, če in samo če velja pogoj: $D_{\mathbf{B}=\mathbf{b},\mathbf{C}_2=\mathbf{c}_2^*}(\omega)^* = 1 \Rightarrow \mathbf{b} = \mathbf{1}$.

Predpostavimo, da je $\mathbf{B}$ edini minimalen zadosten vzrok za D pri ω^* ter da obstaja $\mathbf{b}^* \neq 1$, tako da $D_{\mathbf{B}=\mathbf{b}^*,\mathbf{C}_2=\mathbf{c}_2^*}(\omega)^* = 1$. Definirajmo $\mathbf{B}^\dagger \equiv \mathbf{B}^{[\mathbf{B}=\mathbf{b}^*,\mathbf{C}_2=\mathbf{c}_2^*]}$. Množica spremenljivk $\mathbf{B}^\dagger$ je zadosten vzrok za D pri ω^* velikosti $|\mathbf{C}|$, ki ne vsebuje $\mathbf{B}$, saj je $\mathbf{b}^* \neq 1$. Torej $\mathbf{B} \nsubseteq \mathbf{B}^\dagger$. Če $\mathbf{B}^\dagger$ ni minimalen, potem obstaja $\mathbf{B}' \subset \mathbf{B}^\dagger$, ki je minimalen zadosten vzrok za D pri ω^*. Iz konstrukcije velja $\mathbf{B} \neq \mathbf{B}'$, kar je protislovno z definicijo singularnosti.

Predpostavimo sedaj, da velja $D_{\mathbf{B}=\mathbf{b},\mathbf{C}_2=\mathbf{c}_2^*}(\omega)^* = 1 \Rightarrow \mathbf{b} = \mathbf{1}$ ter da obstaja drug minimalen zadosten vzrok $\mathbf{B}'$ za D pri ω^*, da $\mathbf{B} \neq \mathbf{B}'$. Iz minimalnosti $\mathbf{B}'$ sledi $\mathbf{B} \nsubseteq \mathbf{B}'$. Torej obstaja vrednost $\tilde{\mathbf{c}}$, da $(\mathbf{B})_{\tilde{\mathbf{c}}} \neq 1$, ampak $(\mathbf{B}')_{\tilde{\mathbf{c}}} = 1$. Iz $(\mathbf{B}')_{\tilde{\mathbf{c}}} = 1$ sledi po definiciji zadostnega vzroka $\mathbf{B}'$, da velja $D_{\mathbf{C}=\tilde{\mathbf{c}}}(\omega^*) = 1$, $D_{\mathbf{B}=\tilde{\mathbf{b}},\mathbf{C}_2=\tilde{\mathbf{c}}_2}(\omega^*) = 1$ pri določenih $\tilde{\mathbf{c}}_2$ in $\tilde{\mathbf{b}}$. Po predpostavki ($D_{\mathbf{B}=\mathbf{b},\mathbf{C}_2=\mathbf{c}_2^*}(\omega)^* = 1 \Rightarrow \mathbf{b} = \mathbf{1}$) bi moralo veljati $\tilde{\mathbf{b}} = 1$, vendar $\tilde{\mathbf{b}} = (\mathbf{B})_{\mathbf{B}=\tilde{\mathbf{b}},\mathbf{C}_2=\tilde{\mathbf{c}}_2} = (\mathbf{B})_{\mathbf{C}=\tilde{\mathbf{c}}} \neq 1$. □

Pojma singularnosti in ireducibilnosti ohranjata določene podobnosti, čeprav je singularnost precej strožji pojem. Naj bo $\mathbf{b}_n = (\mathbf{1}_n, 0, \mathbf{1}_{|\mathbf{B}|-n-1})$, kjer je $\mathbf{1}_k$ vektor enic dolžine $k \in \{1, \ldots, |\mathbf{B}| - 1\}$. Ireducibilnost nam zagotavlja le obstoj osebka ω^* in vrednosti $\mathbf{c}_2^*$, da veljata naslednja pogoja za vsak $n \in \{0, \ldots, |\mathbf{B}| - 1\}$:

$$\mathbf{b} = \mathbf{1} \implies D_{\mathbf{B}=\mathbf{b},\mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 1,$$
$$\mathbf{b} = \mathbf{b}_n \implies D_{\mathbf{B}=\mathbf{b},\mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 0.$$

Z naslednjo posledico poudarimo, da je singularen minimalen zadosten vzrok tudi ireducibilen; obratna smer pa ne velja, saj ireducibilna množica ni nujno zadosten vzrok.

Posledica 6.3 *Če je $\boldsymbol{B}$ singularen minimalen zadosten vzrok za $\mathcal{D}(\boldsymbol{C}, \Omega)$, potem je $\boldsymbol{B}$ tudi ireducibilen.*

Dokaz: Sledi iz zgornjega izreka 6.2 ter karakterizacije ireducibilnosti iz izreka 4.2. □

6.2 Singularnost v reprezentaciji zadostnega vzroka

Poglejmo povezavo singularnosti in reprezentacije zadostnega vzroka. Spoznali bomo, da je množica $\mathbf{B}$ singularna, če obstaja vsaj en osebek $\omega^* \in \Omega$, katerega zadosten vzrok je enoličen in vsebuje singularno množico.

Izrek 6.4 *Naj bo $\boldsymbol{B} \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}))$. Množica $\boldsymbol{B}$ je singularna za $\mathcal{D}(\boldsymbol{C}, \Omega)$, če in samo če obstaja $\omega^* \in \Omega$, da v vsaki reprezentaciji $(\boldsymbol{A}, \mathfrak{B})$ za $\mathcal{D}(\boldsymbol{C}, \Omega)$ veljata naslednja pogoja:*

(i) za vse $\boldsymbol{B}^* \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}))$, kjer $|\boldsymbol{B}^*| = |\boldsymbol{C}|$ in $\boldsymbol{B} \subseteq \boldsymbol{B}^*$ obstaja $\boldsymbol{B}_i \in \mathfrak{B}$, tak da $\boldsymbol{B}_i \subseteq \boldsymbol{B}^*$ in velja $A_i(\omega^*) = 1$;

(ii) za vse $\boldsymbol{B}_i \in \mathfrak{B}$, take da $\boldsymbol{B} \nsubseteq \boldsymbol{B}_i$ velja $A_i(\omega^*) = 0$.

Točka (i) zagotavlja, da je $\mathbf{B}$ del zadostnega vzroka za ω^*, točka (ii) pa zagotavlja, da so taki zadostni vzroki, ki vsebujejo $\mathbf{B}$, edini zadostni vzroki za ω^*. Točneje, izrek nam pove, da je $\mathbf{B}$ singularen minimalen vzrok, če (i) je del zadostnega vzroka nekega osebka ω^* in (ii) velja, da tisti zadostni vzroki $\mathbf{B}_j$ iz reprezentacije, ki ne vsebujejo singularne množice $\mathbf{B}$, niso zadostni vzroki za osebek ω^*.

Če je $\mathbf{B}$ minimalen zadostni vzrok, ne more vsebovati zadostnega vzroka $\mathbf{B}_i$, torej iz pogoja (ii) sledi še $\mathbf{B} \subseteq \mathbf{B}_i \subseteq \mathbf{B}^*$.

Pogoj iz posledice 4.6 je podoben pogoju (i) vendar se vsebinsko razlikuje od njega. Zahteva obstoj množice $\mathbf{B}^*$ velikosti $|\mathbf{C}|$, da za vsako reprezentacijo obstaja $\mathbf{B}_j$, da $\mathbf{B} \subseteq \mathbf{B}_i \subseteq \mathbf{B}^*$.

Dokaz: Naj bo $\mathbf{C} = \mathbf{C}_1 \dot\cup \mathbf{C}_2,\ \mathbf{B} \in \dot{\mathbb{P}}(\mathbb{L}(\mathbf{C}_1)),\ |\mathbf{B}| = |\mathbf{C}_1|$.

($\Rightarrow$) Predpostavimo najprej, da je $\mathbf{B}$ singularna za $\mathcal{D}(\mathbf{C}, \Omega)$, torej velja pogoj (6.1) za nek $\omega^* \in \Omega$ ter za vse $\mathbf{c}$. Preverimo, da veljata pogoja (i) in (ii). Naj bo $(\mathbf{A}, \mathfrak{B})$ reprezentacija za $\mathcal{D}(\mathbf{C}, \Omega)$ in naj bo $\mathbf{B}^* \in \mathfrak{B}$, taka da $|\mathbf{B}^*| = |\mathbf{C}|$ in $\mathbf{B} \subseteq \mathbf{B}^*$. Iščemo $\mathbf{B}_i$, da $\mathbf{B} \subseteq \mathbf{B}_i \subseteq \mathbf{B}^*$ in $A_i(\omega^*) = 1$. Uporabimo minimalnost $\mathbf{B}$, ki nam zagotavlja $D_{\mathbf{B}=\mathbf{1},\mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 1$ pri poljubni vrednosti $\mathbf{c}_2^*$. Vzamemo tako vrednost $\mathbf{c}_2^*$, da velja $\mathbf{B}^* = \mathbf{B}^{[\mathbf{B}=\mathbf{1},\mathbf{C}_2=\mathbf{c}_2^*]}$. Iz definicije reprezentacije zadostnega vzroka vemo, da iz $D_{\mathbf{B}=\mathbf{1},\mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 1$ sledi obstoj $A_i \in \mathbf{A},\ \mathbf{B}_i \in \mathfrak{B}$, da $A_i(\omega^*) = 1$ in $(\bigwedge(\mathbf{B}_i))_{\mathbf{B}=\mathbf{1},\mathbf{C}_2=\mathbf{c}_2^*} = 1$. Torej $(\bigwedge(\mathbf{B}_i))_{\mathbf{B}^*=\mathbf{1}} = 1$ in $\mathbf{B}_i \subseteq \mathbf{B}^*$.

Preverili smo pogoj (i) brez uporabe singularnosti. Uporabili ga bomo pri preverjanju pogoja (ii): naj bo $\mathbf{B}_i \in \mathfrak{B}$, taka da $\mathbf{B} \not\subseteq \mathbf{B}_i$. Torej obstaja $\tilde{\mathbf{B}} \neq \mathbf{B}$, da $\mathbf{B}_i \subseteq \mathbf{B}^{[\tilde{\mathbf{B}}=\mathbf{1},\mathbf{C}_2=\tilde{\mathbf{c}}_2]}$, kjer $\tilde{\mathbf{B}} \in \dot{\mathbb{P}}(\mathbb{L}(\mathbf{C}_1))$ in $|\tilde{\mathbf{B}}| = |\mathbf{B}|$. Iz definicije singularnosti $\mathbf{B}$ sledi $D_{\tilde{\mathbf{B}}=\mathbf{1},\mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 0$. Množica $\mathbf{B}_i$ je vsebovana v $\mathbf{B}^{[\tilde{\mathbf{B}}=\mathbf{1},\mathbf{C}_2=\tilde{\mathbf{c}}_2]}$, torej $(\bigwedge(\mathbf{B}_i))_{\tilde{\mathbf{B}}=\mathbf{1},\mathbf{C}_2=\tilde{\mathbf{c}}_2} = 1$ in iz definicije reprezentacije zadostnega vzroka sledi $A_i(\omega^*) = 0$, saj bi v nasprotnem primeru lahko iz $(\bigwedge(\mathbf{B}_i))_{\tilde{\mathbf{B}}=\mathbf{1},\mathbf{C}_2=\tilde{\mathbf{c}}_2} = 1$ sklepali $D_{\tilde{\mathbf{B}}=\mathbf{1},\mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 1$.

($\Leftarrow$) Naj obstaja osebek $\omega^* \in \Omega$, da v vsaki reprezentaciji $(\mathbf{A}, \mathfrak{B})$ velja (i) in (ii). Pokažimo, da velja pogoj (6.1). Desno smer pogoja (6.1) pokažemo s protislovjem: naj velja $D_{\mathbf{B}=\mathbf{b}',\mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 1$ za $\mathbf{b}' \neq \mathbf{1}$. S pomočjo točke (ii) bomo prišli do protislovja. Vzamemo množico $\mathbf{B}' \equiv \mathbf{B}^{[\mathbf{B}=\mathbf{b}']}$ in pogledamo tiste zadostne vzroke $\mathbf{B}_i$ v reprezentaciji, ki vsebujejo $\mathbf{B}'$. Iz pogoja (ii) vemo, da za take $\mathbf{B}_i$ velja $A_i(\omega^*) = 0$, kar je iz definicije reprezentacije ekvivalentno $D_{\mathbf{B}'=\mathbf{1},\mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 0$, saj $(\bigwedge(\mathbf{B}_i))_{\mathbf{B}'=\mathbf{1},\mathbf{C}_2=\mathbf{c}_2^*} = 1$.

Poglejmo še levo smer pogoja (6.1): naj bo $\mathbf{B}^* \equiv \mathbf{B}^{[\mathbf{B}=1,\mathbf{C}_2=\mathbf{c}_2^*]}$ za neko vrednost $\mathbf{c}_2^*$. Iz pogoja (i) sledi obstoj $\mathbf{B}_i \in \mathfrak{B}$, da $\mathbf{B}_i \subseteq \mathbf{B}^*$ in $A_i(\omega^*) = 1$. Iz definicije reprezentacije ter $(\bigwedge(\mathbf{B}_i))_{\mathbf{B}=\mathbf{1},\mathbf{C}_2=\mathbf{c}_2^*} = 1$ sledi $D_{\mathbf{B}'=\mathbf{1},\mathbf{C}_2=\mathbf{c}_2^*}(\omega^*) = 1$. □

6.3 Singularnost pri predpostavki monotonosti

Poglejmo, kaj sledi iz zgornjega izreka.

Posledica 6.5 *Naj bo $\boldsymbol{B} = \boldsymbol{B}_+ \dot\cup \boldsymbol{B}' \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C})),\ |\boldsymbol{B}| = |\boldsymbol{C}|$ in naj $\boldsymbol{B}_+$ pozitivno monotono učinkuje na D glede na $\boldsymbol{C}$. Množica $\boldsymbol{B}$ je singularna za $\mathcal{D}(\boldsymbol{C}, \Omega)$, če in samo če obstaja $\omega^* \in \Omega$, da velja naslednji pogoj:*

$$D_{\boldsymbol{B}=\boldsymbol{1}}(\omega^*) - \sum_{\tilde{\boldsymbol{B}} \subseteq \boldsymbol{B}} D_{\boldsymbol{B} \setminus \tilde{\boldsymbol{B}}=\boldsymbol{1},\tilde{\boldsymbol{B}}=\boldsymbol{0}}(\omega^*) > 0. \tag{6.3}$$

Dokaz: Če velja $|\mathbf{B}| = |\mathbf{C}|$, sta množici $\mathbf{B}^*$ ter $\mathbf{B}_i$ iz pogoja (i) v izreku 6.4 enaki, torej je pogoj (i) ekvivalenten pogoju $A_i(\omega^*) = 1$ za $\mathbf{B}_i = \mathbf{B}^* = \mathbf{B}$. Iz definicije reprezentacije sledi še temu ekvivalenten pogoj $D_{\mathbf{B}_i=\mathbf{1}}(\omega^*) = D_{\mathbf{B}=\mathbf{1}}(\omega^*) = 1$.

Pogoj (ii) v izreku 6.4 zahteva, da za vsako množico $\mathbf{B}_i \in \mathfrak{B}$, ki ne vsebuje $\mathbf{B}$, velja $A_i(\omega^*) = 0$. Iz definicije reprezentacije sledi, da je pogoj (ii) enakovreden zahtevi, da za vsako množico $\mathbf{B}_i \in \mathfrak{B}$, ki ne vsebuje $\mathbf{B}$, velja $D_{\mathbf{B}_i=\mathbf{1}}(\omega^*) = 0$. To pa velja, če in samo če $D_{\mathbf{B}\setminus\tilde{\mathbf{B}}=\mathbf{1},\tilde{\mathbf{B}}=\mathbf{0}}(\omega^*) = 0$ za vse $\tilde{\mathbf{B}} \subseteq \mathbf{B}$. Namreč pri danem $\tilde{\mathbf{B}}$ vzamemo množico

$$\mathbf{B}_i = (\mathbf{B} \setminus \tilde{\mathbf{B}}) \ \cup\ \overline{\tilde{\mathbf{B}}}.$$

$\square$

Poglejmo zadosten pogoj singularnosti pri predpostavki monotonosti, katerega bomo uporabljali pri testiranju singularnosti.

Izrek 6.6 *Naj bo* $\boldsymbol{B} = \boldsymbol{B}_+ \dot{\cup} \boldsymbol{B}' \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}))$, $|\boldsymbol{B}| = |\boldsymbol{C}|$ *in naj* $\boldsymbol{B}_+$ *pozitivno monotono učinkuje na* D *glede na* $\boldsymbol{C}$*. Če za neko drevo* $\mathfrak{T}$ *na* $\boldsymbol{B}_+$ *in nek* $\omega^* \in \Omega$ *velja*

$$D_{\boldsymbol{B}=\boldsymbol{1}}(\omega^*) - \sum_{L \in \boldsymbol{B}_+} D_{\boldsymbol{B}\setminus\{L\}=\boldsymbol{1},L=0}(\omega^*)$$
$$\sum_{\substack{\tilde{\boldsymbol{B}}:\ \tilde{\boldsymbol{B}} \subseteq \boldsymbol{B}' \\ \tilde{\boldsymbol{B}} \neq \emptyset}} D_{\boldsymbol{B}\setminus\tilde{\boldsymbol{B}}=\boldsymbol{1},\tilde{\boldsymbol{B}}=\boldsymbol{0}}(\omega^*) + \sum_{\boldsymbol{E} \in \mathfrak{T}} D_{\boldsymbol{B}\setminus\boldsymbol{E}=\boldsymbol{1},\boldsymbol{E}=\boldsymbol{0}}(\omega^*) > 0, \tag{6.4}$$

potem je $\boldsymbol{B}$ *singularna za* $\mathcal{D}(\boldsymbol{C}, \Omega)$.

Dokaz: Predpostavimo, da velja zgornji pogoj (6.4) za nek $\omega^* \in \Omega$ ter, da ne velja pogoj singularnosti pri predpostavki monotonosti (6.3). Torej za vse $\omega \in \Omega$ velja

$$D_{\mathbf{B}=\mathbf{1}}(\omega) - \sum_{\tilde{\mathbf{B}} \subseteq \mathbf{B}} D_{\mathbf{B}\setminus\tilde{\mathbf{B}}=\mathbf{1},\tilde{\mathbf{B}}=\mathbf{0}}(\omega) \leq 0. \tag{6.5}$$

Zgornjo vsoto bomo najprej omejili na manjšo vsoto s pomočjo pozitivnega monotonega učinka $\mathbf{B}_+$ na D. Pokažimo, da za vse $\omega \in \Omega$ velja naslednja neenačba:

$$D_{\mathbf{B}=\mathbf{1}}(\omega) - \sum_{L \in \mathbf{B}} D_{\mathbf{B}\setminus\{L\}=\mathbf{1},L=0}(\omega) - \sum_{\tilde{\mathbf{B}} \subseteq \mathbf{B}', |\tilde{\mathbf{B}}| \geq 2} D_{\mathbf{B}\setminus\tilde{\mathbf{B}}=\mathbf{1},\tilde{\mathbf{B}}=\mathbf{0}}(\omega) \leq 0. \tag{6.6}$$

Namreč, za vsak osebek ω velja le eden izmed naslednjih pogojev:

(a) $D_{\mathbf{B}=\mathbf{1}}(\omega^*) = 0$,
(b) $D_{\mathbf{B}=\mathbf{1}}(\omega^*) = 1$.

Za osebke ω, pri katerih velja (a), je neenakost (6.6) vsekakor veljavna. Nadalje, za vsak osebek ω velja le eden izmed naslednjih pogojev:

(c) $D_{\mathbf{B}\setminus\{L\}=1,\{L\}=0}(\omega) = 1$ *za nek* $L \in \mathbf{B}$,
(d) $D_{\mathbf{B}\setminus\{L\}=1,\{L\}=0}(\omega) = 0$ *za vse* $L \in \mathbf{B}$.

Če veljata pogoja (b) in (c), potem neenakost 6.6 prav tako velja. Za vsak osebek ω velja le eden izmed naslednjih pogojev:

(e) $D_{\mathbf{B}\setminus\tilde{\mathbf{B}}=\mathbf{1},\tilde{\mathbf{B}}=\mathbf{0}}(\omega) = 1$ *za neko množico* $\tilde{\mathbf{B}} \subset \mathbf{B}'$, $|\tilde{\mathbf{B}}| \geq 2$,
(f) $D_{\mathbf{B}\setminus\tilde{\mathbf{B}}=\mathbf{1},\tilde{\mathbf{B}}=\mathbf{0}}(\omega) = 0$ *za vse množice* $\tilde{\mathbf{B}} \subset \mathbf{B}'$, $|\tilde{\mathbf{B}}| \geq 2$.

Če veljajo pogoji (b), (d) in (e), potem neenakost 6.6 velja. Če pa veljajo pogoji (b), (d) in (f), potem iz enačbe (6.5) sledi obstoj množice $\mathbf{Z}$, kjer $\mathbf{Z} \cap \mathbf{B}_+ \neq \emptyset$, $|\mathbf{Z}| \geq 2$ in $D_{\mathbf{B}\setminus\mathbf{Z}=\mathbf{1},\mathbf{Z}=\mathbf{0}}(\omega) = 1$. Naj bo $Z \in \mathbf{Z} \cap \mathbf{B}_+$. Iz pozitivnega monotonega učinka Z sledi

$$D_{\mathbf{B}\setminus(\mathbf{Z}\setminus\{Z\})=\mathbf{1},\mathbf{Z}\setminus\{Z\}=\mathbf{0}}(\omega) = 1, \tag{6.7}$$

saj $D_{\mathbf{B}\setminus(\mathbf{Z}\setminus\{Z\})=\mathbf{1},\mathbf{Z}\setminus\{Z\}=\mathbf{0}}(\omega) \geq D_{\mathbf{B}\setminus\mathbf{Z}=\mathbf{1},\tilde{\mathbf{B}}=\mathbf{0}}(\omega) = 1$. Velja lahko samo eden izmed naslednjih pogojev:

(g) $|\mathbf{Z} \setminus \{Z\}| = 1$,
(h) $|\mathbf{Z} \setminus \{Z\}| \geq 2$.

Če veljajo pogoji (b), (d), (f) in (g), potem je enakost (6.7) v nasprotju s predpostavko (d). Torej ne morejo obenem veljati pogoji (b), (d), (f) in (g). Velja lahko samo eden izmed naslednjih pogojev:

$$(i)\ \mathbf{Z} \setminus \{Z\} \subset \mathbf{B}'$$
$$(j)\ \mathbf{Z} \setminus \{Z\} \cap \mathbf{B}_+ \neq \emptyset.$$

Če veljajo pogoji (b), (d), (f), (h) in (i), potem je enakost (6.7) v nasprotju s predpostavko (f). Če pa veljajo pogoji (b), (d), (f), (h) in (j), potem iz enačbe (6.5) sledi obstoj množice $\mathbf{Z}_2$, kjer $\mathbf{Z}_2 \cap \mathbf{B}_+ \neq \emptyset$ in $|\mathbf{Z}_2| \geq 2$ in $D_{\mathbf{B}\setminus\mathbf{Z}_2=\mathbf{1},\mathbf{Z}_2=\mathbf{0}}(\omega) = 1$. Nadaljujemo kot pri pogoju (f), dokler ne dobimo množice $\mathbf{Z}_n$, da $|\mathbf{Z}_n \setminus Z_n| = 1$, $\mathbf{Z}_n \cap \mathbf{B}_+ \neq \emptyset$ in $D_{\mathbf{B}\setminus\mathbf{Z}_n=\mathbf{1},\mathbf{Z}_n=\mathbf{0}}(\omega) = 1$, kar je v nasprotju s predpostavko (d). Zaključimo, da ne morejo obenem veljati pogoji (b), (d) in (f), torej velja enačba (6.6).

Podobno kot v dokazu izreka 4.14 iz pogoja (6.6) sledi, da za vsak ω obstaja $R \in \mathbf{B}_+$, da velja

$$D_{\mathbf{B}=\mathbf{1}}(\omega) - \sum_{L\in\mathbf{B}'\cup\{R\}} D_{\mathbf{B}\setminus\{L\}=\mathbf{1},L=0}(\omega) - \sum_{\tilde{\mathbf{B}}\subseteq\mathbf{B}',|\tilde{\mathbf{B}}|\geq 2} D_{\mathbf{B}\setminus\tilde{\mathbf{B}}=\mathbf{1},\tilde{\mathbf{B}}=\mathbf{0}}(\omega) \leq 0. \tag{6.8}$$

Naj bo $\mathfrak{T}$ drevo na B_+. Izpišimo levo stran enačbe (6.4) tako, da vsote razstavimo:

$$\begin{aligned}
&D_{\mathbf{B}=\mathbf{1}}(\omega) - \sum_{L\in\mathbf{B}_+\setminus\{R\}} D_{\mathbf{B}\setminus L=\mathbf{1},L=0}(\omega) - D_{\mathbf{B}\setminus\{R\}=\mathbf{1},R=0}(\omega) - \sum_{\tilde{\mathbf{B}}\subseteq\mathbf{B}',\ |\tilde{\mathbf{B}}|\geq 2} D_{\mathbf{B}\setminus\tilde{\mathbf{B}}=\mathbf{1},\tilde{\mathbf{B}}=\mathbf{0}}(\omega) \\
&\qquad - \sum_{L\in\mathbf{B}'} D_{\mathbf{B}\setminus L=\mathbf{1},L=0}(\omega) + \sum_{\mathbf{E}\in\mathfrak{T}} D_{\mathbf{B}\setminus\mathbf{E}=\mathbf{1},\mathbf{E}=\mathbf{0}}(\omega) \\
&= \left(D_{\mathbf{B}=\mathbf{1}}(\omega) - \sum_{L\in\mathbf{B}'\cup\{R\}} D_{\mathbf{B}\setminus\{L\}=\mathbf{1},L=0}(\omega) - \sum_{\tilde{\mathbf{B}}\subseteq\mathbf{B}',\ |\tilde{\mathbf{B}}|\geq 2} D_{\mathbf{B}\setminus\tilde{\mathbf{B}}=\mathbf{1},\tilde{\mathbf{B}}=\mathbf{0}}(\omega)\right) \\
&\qquad + \left(\sum_{L\in\mathbf{B}_+\setminus\{R\}} -D_{\mathbf{B}\setminus L=\mathbf{1},L=0}(\omega) + D_{\mathbf{B}\setminus\phi_{\mathbf{R}}^{\mathfrak{T}}(L)=\mathbf{1},\ \phi_{\mathbf{R}}^{\mathfrak{T}}(L)=\mathbf{0}}(\omega)\right),
\end{aligned}$$

kjer je $\phi_{\mathbf{R}}^{\mathfrak{T}}$ bijekcija iz trditve 4.12 in $\phi_{\mathbf{R}}^{\mathfrak{T}}(L) = \mathbf{E} = \{L, L'\}$ zadnja povezava na poti iz R v L. Če upoštevamo neenačbo (6.8) ter pozitiven monoton učinek L' na D, sledi, da je zgornji izraz negativen, oziroma enak 0. Prišli smo v protislovje s predpostavko. $\square$

Poglejmo, kaj sledi iz izreka 6.6. Spoznali bomo, da ima predpostavka monotonosti velik vpliv na pogoj singularnosti. Namreč, če večina dogodkov iz $\mathbf{B}$ pozitivno monotono učinkuje na D in je $|\mathbf{B}| = |\mathbf{C}|$, potem je singularnost ekvivalentna ireducibilnosti:

Trditev 6.7 *Naj bo $|\boldsymbol{B}| = |\boldsymbol{C}|$. Naj vsak (ali vsak razen enega) izmed $B_i \in \boldsymbol{B}$ pozitivno monotono učinkuje na D glede na $\boldsymbol{B}$. Množica $\boldsymbol{B}$ je singularna za $\mathcal{D}(\boldsymbol{C}, \Omega)$, če in samo če je $\boldsymbol{B}$ ireducibilna za $\mathcal{D}(\boldsymbol{C}, \Omega)$.*

Dokaz: Poglejmo najprej levo smer dokaza. Predpostavimo, da je $\mathbf{B}$ ireducibilna in dokažimo, da je singularna s pomočjo zadostnega pogoja iz izreka 6.6. Ta se v primeru, ko je množica $\mathbf{B}'$ velikosti $|\mathbf{B}'| \leq 1$, glasi

$$D_{\mathbf{B}=\mathbf{1}}(\omega^*) - \sum_{L\in\mathbf{B}} D_{\mathbf{B}\setminus\{L\}=\mathbf{1},L=0}(\omega^*) + \sum_{\mathbf{E}\in\mathfrak{T}} D_{\mathbf{B}\setminus\mathbf{E}=\mathbf{1},\mathbf{E}=\mathbf{0}}(\omega^*) > 0$$

za nek osebek ω^* ter drevo $\mathfrak{T}$. Zgornji pogoj velja pri ireducibilnih množicah, saj smo v dokazu izreka 4.14 spoznali, da če velja obratna neenakost za vsak ω^*, potem $\mathbf{B}$ ni ireducibilna.

Desna smer dokaza sledi iz posledice 6.3. $\square$

6.4 Ohranjanje singularnosti pri večjih množicah

Poglejmo, kdaj je singularna množica za $\mathcal{D}(\mathbf{C}, \Omega)$, singularna tudi za $\mathcal{D}(\mathbf{C} \cup \mathbf{C}', \Omega)$.

Trditev 6.8 *Naj bo $|\boldsymbol{B}| = |\boldsymbol{C}|$ in naj vsak (ali vsak razen enega) izmed $B_i \in \boldsymbol{B} \dot{\cup} \boldsymbol{C'}$ pozitivno monotono učinkuje na D glede na $\boldsymbol{B} \cup \boldsymbol{C'}$. Predpostavimo še, da množica $\boldsymbol{C}$ ne vpliva kavzalno na $\boldsymbol{C'}$. Če je $\boldsymbol{B}$ singularen za $\mathcal{D}(\boldsymbol{C}, \Omega)$, potem je $\boldsymbol{B}$ singularen za $\mathcal{D}(\boldsymbol{C} \cup \boldsymbol{C'}, \Omega)$.*

Dokaz: Iz trditve 6.7 sledi, da je $\mathbf{B}$ ireducibilen glede na $\mathcal{D}(\mathbf{C}, \Omega)$. Po trditvi 4.8 je $\mathbf{B}$ ireducibilen glede na $\mathcal{D}(\mathbf{C} \cup \mathbf{C}', \Omega)$. Ponovno iz trditve 6.7 sledi še, da je $\mathbf{B}$ singularen za $\mathcal{D}(\mathbf{C} \cup \mathbf{C}', \Omega)$. □

6.5 Testi singularnosti

Obravnavali bomo teste singularnosti pri predpostavki monotonosti, ki neposredno sledijo iz izreka 6.6. Testi veljajo tudi v posebnem primeru, ko noben dogodek iz množice $\mathbf{B}$ ne učinkuje pozitivno monotono.

Posledica 6.9 *Naj bo $\boldsymbol{B} = \boldsymbol{B}_+ \dot{\cup} \boldsymbol{B}' \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}))$, $|\boldsymbol{B}| = |\boldsymbol{C}|$ in naj $\boldsymbol{B}_+$ pozitivno monotono učinkuje na D glede na $\boldsymbol{B}$. Naj $\boldsymbol{W}$ zadošča pri popravljanju zavajajočega učinka $\boldsymbol{C}$ na D. Če za neko drevo $\mathfrak{T}$ na $\boldsymbol{B}_+$ in neko vrednost $\boldsymbol{w}$ velja*

$$
\begin{aligned}
&E[D \mid \boldsymbol{B} = \boldsymbol{1}, \boldsymbol{W} = \boldsymbol{w}] - \sum_{L \in \boldsymbol{B}_+} E[D \mid \boldsymbol{B} \setminus \{L\} = \boldsymbol{1}, L = 0, \boldsymbol{W} = \boldsymbol{w}] \\
&- \sum_{\substack{\tilde{\boldsymbol{B}}:\ \tilde{\boldsymbol{B}} \subseteq \boldsymbol{B}' \\ \tilde{\boldsymbol{B}} \neq \emptyset}} E[D \mid \boldsymbol{B} \setminus \tilde{\boldsymbol{B}} = \boldsymbol{1}, \tilde{\boldsymbol{B}} = \boldsymbol{0}, , \boldsymbol{W} = \boldsymbol{w}] \\
&+ \sum_{\boldsymbol{E} \in \mathfrak{T}} E[D \mid \boldsymbol{B} \setminus \boldsymbol{E} = \boldsymbol{1}, \boldsymbol{E} = \boldsymbol{0}, \boldsymbol{W} = \boldsymbol{w}] > 0,
\end{aligned}
\tag{6.9}
$$

potem je $\boldsymbol{B}$ singularna za $\mathcal{D}(\boldsymbol{C}, \Omega)$.

Dokaz: Uporabimo trditev 5.2 pri vsakem členu iz enačbe (6.4). □

V primeru, da ni monotonega učinka, izpadeta drugi in četrti člen leve strani neenačbe 6.9. Dobimo pogoj singularnosti brez predpostavke monotonosti, ki je zahtevnejši od pogoja pri predpostavki monotonosti in ga lahko uporabimo kot test prisotnosti kompozicijske epistaze.

Če $|\mathbf{B}| = |\mathbf{C}|$ in vsak (ali vsak razen enega) izmed $B_i \in \mathbf{B}$ pozitivno monotono učinkuje na D glede na $\mathbf{B}$, potem preverjamo singularnost s pomočjo testov ireducibilnosti iz izreka 5.4. Zadnje sledi iz trditve 6.7.

7 Interakcije v regresijskem modelu

V naslednjem poglavju bomo obravnavali pogoje ireducibilnosti, minimalnosti in singularnosti v nasičenem Bernoullijevem regresijskem modelu. Spoznali bomo, da se razlikujejo od klasičnih pogojev, ki nam opisujejo prisotnost interakcij med spremenljivkami v regresijskem modelu.

Obravnavajmo najprej Bernoullijeve linearne regresijske modele in nasičene modele. Spoznali bomo, da se nasičeni modeli ujemajo s funkcijo pogojnega matematičnega upanja. Nato bomo obravnavali pogoje ireducibilnosti, minimalnosti in singularnosti v povezavi z regresijskimi koeficienti v teh modelih.

7.1 Bernoullijev linearen regresijski model

Linearen regresijski model povzema linearno povezavo med spremenljivkami, ki predstavljajo vzroke in spremenljivko, ki predstavlja posledico. To je linearna enačba

$$D = \alpha_1 Z_1 + \alpha_2 Z_2 + \cdots + \alpha_s Z_s + \epsilon,$$

kjer so slučajne spremenljivke $Z_1, \ldots, Z_s, D, \epsilon$ definirane na populaciji Ω ter so $\alpha_1, \ldots, \alpha_s$ fiksni koeficienti. Slučajnim spremenljivkam $Z_1, \ldots, Z_s$ pravimo *pojasnjevalne spremenljivke* (explanatory variables) ali *regresorji*, neznane parametre $\alpha_1, \ldots, \alpha_s$ imenujemo *regresijski koeficienti*, slučajno spremenljivko D imenujemo *odvisna spremenljivka*, slučajno spremenljivko ϵ pa *napaka*. Regresijski koeficient α_i določa vpliv spremenljivke Z_i na spremenljivko D.

Bernoullijev regresijski model je regresijski model, v katerem je spremenljivka D binarna in zavzema le vrednosti 0 ali pa 1. Obravnavali bomo Bernoullijeve regresijske modele, v katerih so tudi pojasnjevalne spremenljivke binarne.

7.2 Nasičen model

Nasičen regresijski model je tak model, ki vsebuje koeficiente za vse možne izide, oziroma upošteva vse interakcije med spremenljivkami. *Interakcijo* med dvema slučajnima spremenljivkama X_1 in X_2 definiramo kot produkt $X_1 X_2$. Interakcije med slučajnimi spremenljivkami $X_1, \ldots X_p$ pa so produkti $\prod_{j=1}^{m} Y_j$, kjer je $\{Y_1, \ldots, Y_m\}$ podmnožica $\{X_1, \ldots, X_p\}$.

Natančneje, linearen regresijski model z binarnimi pojasnjevalnimi spremenljivkami je *nasičen*, če za vsako interakcijo med pojasnjevalnimi spremenljivkami, obstaja regresijski koeficient.

Naj bodo $X_1, \ldots, X_p$ pojasnjevalne spremenljivke ter naj bo $(Y_1, \ldots, Y_m)$ vektor pojasnjevalnih spremenljivk, kjer je $\{Y_1, \ldots, Y_m\}$ podmnožica $\{X_1, \ldots, X_p\}$. Naj bo $(y_1, \ldots, y_m) \in \{0,1\}^m$. Če je model nasičen, potem za vsak $(y_1, \ldots, y_m) \in \{0,1\}^m$ obstaja koeficient $\alpha_{y_1,\ldots,y_m}$, da velja

$$E[D|Y_1 = y_1, \ldots, Y_m = y_m] = \alpha_{y_1,\ldots,y_m}.$$

Nasičen linearen regresijski model je torej naslednje oblike:

$$D = \sum_{\{Y_1,\ldots,Y_s\} \subseteq \{X_1,\ldots,X_p\}} \beta_{\{Y_1,\ldots,Y_s\}} \prod_{j=1}^{s} Y_j,$$

kjer so $\beta_{\{Y_1,\ldots,Y_s\}}$ regresijski koeficienti. Koeficientom $\beta_{\{Y_i\}}$ pravimo *glavni učinki* (main effects), ostalim regresijskim koeficientom pa *interakcijski členi*.

Zgornji model je linearen v spremenljivkah $\{Z_k\}_{k \in \{1,\ldots,m\}}$, kjer je m število podmnožic množice $\{X_1, \ldots, X_p\}$, $Z_i = \prod_{j=1}^{m} Y_j$ ter $\{Y_1, \ldots, Y_m\} \subseteq \{X_1, \ldots, X_p\}$. Torej se model ujema s funkcijo pogojnega matematičnega upanja, ki je tudi linearna:

$$E[D \mid X_1 = x_1, \ldots, X_p = x_p] = \sum_{\{Y_1,\ldots,Y_m\} \subseteq \{X_1,\ldots,X_p\}} \beta_{(Y_1,\ldots,Y_m)} \prod_{j=1}^{m} y_j,$$

kjer je $(y_1, ..., y_m) \in \{0,1\}^s$ vrednost vektorja $(Y_1, \ldots, Y_m)$ v primeru, da zavzema vektor $(X_1, \ldots, X_p)$ vrednost $(x_1, \ldots, x_p) \in \{0,1\}^m$.

Če označimo $\mathbf{C} = \{X_1, \ldots, X_p\}$ in $\mathbf{c} = \{x_1, \ldots, x_p\}$ lahko zapišemo linearen nasičen regresijski model na sledeči način:

$$E[D \mid \mathbf{C} = \mathbf{c}] = \sum_{\tilde{\mathbf{B}} \subseteq \mathbf{C}} \beta_{\tilde{B}} \left(\bigwedge (\tilde{\mathbf{B}}) \right)_{\mathbf{c}}. \tag{7.1}$$

Primer 7.1 *Poglejmo primer linearnega nasičenega regresijskega modela z binarno spremenljivko D ter dvema binarnima pojasnjevalnima spremenljivkama X_1 in X_2. Naj X_1 označuje spol, oziroma*

$X_1(\omega) = 1$, če je ω ženska, ter $X_1(\omega) = 0$, če je ω moški. Naj X_2 označuje določeno bolezen, oziroma $X_2(\omega) = 1$, če je osebek ω bolan, ter $X_2(\omega) = 0$, če je osebek ω zdrav. Nasičen model zapišemo na sledeči način:

$$D = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{12} X_1 X_2,$$

kjer označimo $\beta_0 = \beta_\emptyset$, $\beta_1 = \beta_{\{X_1\}}$, $\beta_2 = \beta_{\{X_2\}}$, $\beta_{12} = \beta_{\{X_1,X_2\}}$. Zapis modela ni enoličen. Lahko ga zapišemo tudi s spremenljivkami, Z_1, Z_2, Z_3 ter Z_4, kjer $Z_1(\omega) = 1$, če je ω bolna ženska, $Z_1(\omega) = 0$, če pa ω ni bolna ženska, podobno Z_2 označuje zdravo žensko, Z_3 bolnega moškega in Z_4 zdravega moškega. V tem primeru zapišemo regresijski model v naslednji obliki:

$$D = \alpha_1 Z_1 + \alpha_2 Z_2 + \alpha_3 Z_3 + \alpha_4 Z_4.$$

V tem zapisu modela jasno vidimo, da nasičen regresijski model vsebuje koeficiente za vse možne izide. Koeficiente zgornjih dveh parametrizacij modela povezujejo naslednje enačbe:

$$E[D \mid X_1 = 0, X_2 = 0] = E[D \mid Z_4 = 0] = \beta_0 = \alpha_1,$$

$$E[D \mid X_1 = 1, X_2 = 0] = E[D \mid Z_2 = 0] = \beta_0 + \beta_1 = \alpha_2,$$
$$E[D \mid X_1 = 0, X_2 = 1] = E[D \mid Z_3 = 0] = \beta_0 + \beta_2 = \alpha_3,$$
$$E[D \mid X_1 = 1, X_2 = 1] = E[D \mid Z_1 = 0] = \beta_0 + \beta_1 + \beta_2 + \beta_{12} = \alpha_4.$$

7.3 Interakcije v regresijskem modelu

Zanimajo nas pogoji ireducibilnosti, minimalnosti in singularnosti izraženi z regresijskimi koeficienti. Za motivacijo poglejmo najprej preprost primer z dvema pojasnjevalnima spremenljivkama

Primer 7.2 *Naj bosta X_1 in X_2 dogodka, ki vplivata na binarni izid D. Predpostavimo, da X_1 in X_2 ne učinkujeta zavajajoče na D, oziroma velja*

$$D_{X_1=x_1, X_2=x_2} \perp\!\!\!\perp (X_1, X_2)$$

za vse x_1, x_2. Označimo

$$p_{x_1,x_2} = P[D = 1 \mid X_1 = x_1, X_2 = x_2] = E[D \mid X_1 = x_1, X_2 = x_2].$$

Test ireducibilnosti brez pogoja monotonosti iz izreka 5.3 je enak naslednjemu pogoju:

$$p_{11} - p_{10} - p_{01} > 0.$$

Test ireducibilnosti pri pogoju monotonosti iz izreka 5.14 pa naslednjemu pogoju:

$$p_{11} - p_{10} - p_{01} + p_{00} > 0.$$

Obravnavajmo še nasičen linearen regresijski model:

$$E[D \mid X_1 = x_1, X_2 = x_2] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2.$$

V klasični literaturi, ki obravnava regresijo, imamo poenostavljen model, ki pravi, da poteka interakcija med členoma X_1 in X_2, če velja $\beta_{12} > 0$. Izpišimo zgornje pogoje ireducibilnosti z uporabo regresijskih koeficientov. V ta namen poglejmo povezavo med regresijskimi koeficienti in p_{x_1,x_2} podano v naslednjih enačbah:

$$p_{11} = E[D \mid X_1 = 1, X_2 = 1] = \beta_0 + \beta_1 + \beta_2 + \beta_{12},$$

$$p_{10} = E[D \mid X_1 = 1, X_2 = 0] = \beta_0 + \beta_1,$$
$$p_{01} = E[D \mid X_1 = 0, X_2 = 1] = \beta_0 + \beta_2,$$
$$p_{00} = E[D \mid X_1 = 0, X_2 = 0] = \beta_0.$$

Sledi pogoj ireducibilnosti brez predpostavke monotonosti iz izreka 5.3 izražen z regresijskimi koeficienti,

$$\beta_{12} > \beta_0,$$

ter pogoj ireducibilnosti pri predpostavki monotonosti iz izreka 5.14 izražen z regresijskimi koeficienti,

$$\beta_{12} > 0.$$

Opazimo, da je v tem primeru pojem interakcije iz klasične literature enakovreden le pojmu ireducibilnosti pri predpostavki monotonosti. Razlikuje pa se s pojmom ireducibilonosti brez predpostavke monotonosti.

7.3.1 Pogoji ireducibilnosti

V sledečih poglavjih bomo posplošili zgornji primer, oziroma spoznali splošne pogoje ireducibilnosti izražene z regresijskimi koeficienti. Najprej bomo podali zadosten pogoj ireducibilnosti brez predpostavke monotonosti, ki je ekvivalenten pogoju iz izreka 5.3. Pogoj bo oblike $\sum_{\tilde{\mathbf{B}} \subset \mathbf{C}} m_{\tilde{B}} \beta_{\tilde{B}} > 0$, kjer bodo $m_{\tilde{B}}$ točno določeni koeficienti.

Trditev 7.1 *Naj bo $\boldsymbol{C} = \boldsymbol{B} \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}))$. Naslednja pogoja sta ekvivalentna:*

$$(i)\ E[D \mid \boldsymbol{B} = \boldsymbol{1}] - \sum_{L \in \boldsymbol{B}} E[D \mid \boldsymbol{B} \setminus \{L\} = \boldsymbol{1},\ L = 0] > 0,$$
$$(ii) \sum_{\tilde{\boldsymbol{B}} \subseteq \boldsymbol{C}} (1 - |\boldsymbol{B} \setminus \tilde{\boldsymbol{B}}|)\ \beta_{\tilde{\boldsymbol{B}}} > 0.$$

Torej je zadosten pogoj ireducibilnosti v regresijskem modelu enak pogoju (ii). Poudarimo, da vsota iz točke (ii) vključuje konstanto $\beta_\emptyset$.

Dokaz: Iz Bernoullijevega regresijskega modela velja

$$E[D \mid \mathbf{B} = \mathbf{1}] = \sum_{\tilde{\mathbf{B}} \subseteq \mathbf{B}} \beta_{\tilde{B}} \left(\bigwedge(\tilde{\mathbf{B}}) \right)_{\mathbf{B}=\mathbf{1}} = \sum_{\tilde{\mathbf{B}} \subseteq \mathbf{B}} \beta_{\tilde{\mathbf{B}}}$$

ter

$$\sum_{L \in \mathbf{B}} E[D \mid \mathbf{B} \setminus \{L\} = \mathbf{1},\ L = 0] = \sum_{L \in \mathbf{B}} \sum_{\tilde{\mathbf{B}} \subseteq \mathbf{B}} \beta_{\tilde{B}} \left(\bigwedge(\tilde{\mathbf{B}}) \right)_{\mathbf{B} \setminus \{L\} = \mathbf{1},\ L=0}$$
$$= \sum_{L \in \mathbf{B}} \sum_{\tilde{\mathbf{B}} \subseteq \mathbf{B} \setminus \{L\}} \beta_{\tilde{B}},$$

saj velja enakost $(\bigwedge(\tilde{\mathbf{B}}))_{\mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0} = 0$, če je dogodek L podmnožica $\tilde{\mathbf{B}}$ ter velja enakost $(\bigwedge(\tilde{\mathbf{B}}))_{\mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0} = 1$, če dogodek L ni podmnožica $\tilde{\mathbf{B}}$. Če je množica $\tilde{\mathbf{B}}$ sestavljena iz enega elementa je število ponavljanj koeficienta $\beta_{\tilde{\mathbf{B}}}$ v zgornji dvojni vsoti enako $(|\mathbf{B}| - 1)$; če je množica $\tilde{\mathbf{B}}$ sestavljena iz dveh elementov je število ponavljanj koeficienta $\beta_{\tilde{\mathbf{B}}}$ v zgornji dvojni vsoti enako

($|\mathbf{B}|-2$) itd. Zaključimo brez formalnega preverjanja, da je število ponavljanj koeficienta $\beta_{\tilde{\mathbf{B}}}$ v zgornji dvojni vsoti enako $|\mathbf{B}\setminus\tilde{\mathbf{B}}|$. Torej velja

$$\sum_{L\in\mathbf{B}} E[D \mid \mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0] = \sum_{\tilde{\mathbf{B}}\subseteq\mathbf{C}} |\mathbf{B}\setminus\tilde{\mathbf{B}}|\ \beta_{\tilde{\mathbf{B}}}.$$

Sledi, da je pogoj (ii) ekvivalenten pogoju

$$E[D \mid \mathbf{B}=\mathbf{1}] - \sum_{L\in\mathbf{B}} E[D \mid \mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0] > 0 \tag{7.2}$$

□

7.3.2 Pogoji ireducibilnosti pri predpostavki monotonosti

Poglejmo, kateri pogoji regresijskih koeficientov so zadostni pogoji ireducibilnosti pri predpostavki monotonosti. V naslednji trditvi bomo spoznali pogoj, ki je ekvivalenten pogoju iz izreka 4.14. Najprej definirajmo *stopnjo dogodka* L v drevesu $\mathfrak{T}$ kot število povezav v $\mathfrak{T}$, ki vsebujejo L,

$$deg_{\mathfrak{T}}(L) \equiv \mid \{\mathbf{E} | \mathbf{E}\in\mathfrak{T}, L\in\mathfrak{E}\}|,$$

ter *stopnjo množice dogodkov* $\tilde{\mathbf{B}}$ v drevesu $\mathfrak{T}$,

$$deg_{\mathfrak{T}}(\tilde{\mathbf{B}}) \equiv \sum_{L\in\tilde{\mathbf{B}}} deg_{\mathfrak{T}}(L).$$

Označimo $\mathfrak{T}_{\tilde{\mathbf{B}}}$ podgraf $\mathfrak{T}$ -ja, ki ga inducira podmnožica $\tilde{\mathbf{B}}$:

$$\mathfrak{T}_{\tilde{\mathbf{B}}} = \{\mathbf{E} \mid \mathbf{E}\in\mathfrak{T}, \mathbf{E}\subseteq\tilde{\mathbf{B}}\cap\mathbf{B}_+\}.$$

Trditev 7.2 *Naj bo $\boldsymbol{C}=\boldsymbol{B}=(\boldsymbol{B}_+\dot\cup\boldsymbol{B}')\in\dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}))$ in naj $\boldsymbol{B}_+$ pozitivno monotono učinkuje na D glede na $\boldsymbol{C}$. Naslednja pogoja za neko drevo $\mathfrak{T}$ na $\boldsymbol{B}_+$ sta ekvivalentna:*

$$\begin{aligned}
&(i) E[D \mid \boldsymbol{B}=\boldsymbol{1}] - \sum_{L\in\boldsymbol{B}} E[D \mid \boldsymbol{B}\setminus\{L\}=\boldsymbol{1},\ L=0] \\
&\quad + \sum_{\boldsymbol{E}\in\mathfrak{T}} E[D \mid \boldsymbol{B}\setminus\boldsymbol{E}=\boldsymbol{1},\ \boldsymbol{E}=\boldsymbol{0}] > 0. \\
&(ii) \sum_{\tilde{\boldsymbol{B}}\subseteq\boldsymbol{C}} m^{ired}_{\tilde{\boldsymbol{B}}}\beta_{\tilde{\mathbf{B}}} > 0,\ \textit{kjer} \\
&\quad m^{ired}_{\tilde{\boldsymbol{B}}} = 1 - |\boldsymbol{B}\setminus\tilde{\boldsymbol{B}}| + |\mathfrak{T}| - deg_{\mathfrak{T}}(\tilde{\mathbf{B}}) + |\mathfrak{T}_{\tilde{\mathbf{B}}}|.
\end{aligned}$$

Dokaz: Poglejmo najprej naslednji izraz:

$$\begin{aligned}
\sum_{\mathbf{E}\in\mathfrak{T}} E[D \mid \mathbf{B}\setminus\mathbf{E}=\mathbf{1},\ \mathbf{E}=\mathbf{0}] &= \sum_{\mathbf{E}\in\mathfrak{T}}\ \sum_{\tilde{\mathbf{B}}\subseteq\mathbf{B}} \beta_{\tilde{B}}\Big(\bigwedge(\tilde{\mathbf{B}})\Big)_{\mathbf{B}\setminus\mathbf{E}=\mathbf{1},\ \mathbf{E}=\mathbf{0}} \\
&= \sum_{\mathbf{E}\in\mathfrak{T}}\ \sum_{\tilde{\mathbf{B}}\subseteq\mathbf{B}\setminus\mathbf{E}} \beta_{\tilde{B}}.
\end{aligned}$$

Upoštevali smo, da velja enakost $(\bigwedge(\tilde{\mathbf{B}}))_{\mathbf{B}\setminus\mathbf{E}=\mathbf{1},\ \mathbf{E}=\mathbf{0}} = 1$ le v primeru, ko povezava $\mathbf{E}$ ni vsebovana v $\tilde{\mathbf{B}}$. Če je množica $\tilde{\mathbf{B}}$ sestavljena iz enega elementa je število ponavljanj koeficienta $\beta_{\tilde{\mathbf{B}}}$ v zgornji dvojni vsoti enako $|\mathfrak{T}| - deg_{\mathfrak{T}}(\tilde{\mathbf{B}})$. V splošnem pa je število pojavitev koeficienta $\beta_{\tilde{\mathbf{B}}}$ enako

$$|\mathfrak{T}| - \sum_{\mathbf{E}\in\mathfrak{T},\mathbf{E}\cap\tilde{\mathbf{B}}=\emptyset} 1.$$

Zdi se, da je vsota tistih povezav, ki sekajo množico $\tilde{\mathbf{B}}$, $\sum_{\mathbf{E}\in\mathfrak{T},\mathbf{E}\cap\tilde{\mathbf{B}}=\emptyset} 1$, enaka stopnji $deg_{\mathfrak{T}}(\tilde{\mathbf{B}})$, vendar ni tako. Upoštevati moramo še število tistih povezav, ki smo jih dvakrat prešteli. Namreč iz definiciji stopnje sledi, da moramo vsako povezavo, ki je vsebovana v $\tilde{\mathbf{B}}$ prešteti dvakrat: pri vsakem krajišču povezave jo preštejemo enkrat. Torej moramo stopnji odšteti še število povezav v $\tilde{\mathbf{B}}$. Število ponavljanj koeficienta $\beta_{\tilde{\mathbf{B}}}$ v zgornji dvojni vsoti je torej enako

$$|\mathfrak{T}| - (deg_{\mathfrak{T}}(\tilde{\mathbf{B}}) - |\mathfrak{T}_{\tilde{\mathbf{B}}}|).$$

Sledi:

$$\sum_{\mathbf{E}\in\mathfrak{T}} E[D \mid \mathbf{B}\setminus\mathbf{E}=\mathbf{1},\ \mathbf{E}=\mathbf{0}] = \sum_{\tilde{\mathbf{B}}\subseteq\mathbf{B}} \Big(|\mathfrak{T}| - deg_{\mathfrak{T}}(\tilde{\mathbf{B}}) + |\mathfrak{T}_{\tilde{\mathbf{B}}}|\Big)\beta_{\tilde{B}}. \tag{7.3}$$

Iz trditve 7.1 sledi, da je pogoj (ii) ekvivalenten pogoju

$$\begin{aligned} E[D \mid \mathbf{B}=\mathbf{1}] &- \sum_{L\in\mathbf{B}} E[D \mid \mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0] \\ &+ \sum_{\mathbf{E}\in\mathfrak{T}} E[D \mid \mathbf{B}\setminus\mathbf{E}=\mathbf{1},\ \mathbf{E}=\mathbf{0}] > 0. \end{aligned}$$

□

Dokazali smo, da je pogoj $\sum_{\tilde{\mathbf{B}}\subseteq\mathbf{C}} m_{\tilde{\mathbf{B}}}^{irred}\beta_{\tilde{\mathbf{B}}} > 0$ zadosten pogoj ireducibilnosti pri predpostavki monotonosti v Bernoullijevem regresijskem modelu.

7.3.3 Pogoji minimalnosti pri predpostavki monotonosti

Poglejmo še pogoje regresijskih koeficientov, ki ustrezajo pogojem minimalnosti zadostnih vzrokov iz izrekov 5.5 in 5.6.

Trditev 7.3 *Naj bo $\boldsymbol{C} = \boldsymbol{C}_1\dot\cup\boldsymbol{C}_2$, $\boldsymbol{B} \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}_1))$, $|\boldsymbol{B}| = |\boldsymbol{C}_1|$. Naj $\boldsymbol{C}_2$ pozitivno monotono učinkuje na D glede na $\boldsymbol{C}$. Naslednja pogoja sta ekvivalentna:*

$$(i) E[D \mid \boldsymbol{B}=\boldsymbol{1}, \boldsymbol{C}_2=\boldsymbol{0}] - \sum_{L\in\boldsymbol{B}} E[D \mid \boldsymbol{B}\setminus\{L\}=\boldsymbol{1},\ L=0,\ \boldsymbol{C}_2=\boldsymbol{0}] > 0,$$

$$(ii) \sum_{\tilde{\boldsymbol{B}}\subseteq\boldsymbol{C}} (1 - |\boldsymbol{B}\setminus\tilde{\boldsymbol{B}}|)\ \beta_{\tilde{\boldsymbol{B}}} > 0.$$

Dokaz: V nasičenem linearnem Bernoullijevem regresijskem modelu veljajo naslednje neenakosti:

$$E[D \mid \mathbf{B}=\mathbf{1}, \mathbf{C}_2=\mathbf{0}] = \sum_{\tilde{\mathbf{B}}\subseteq\mathbf{B}\cup\mathbf{C}_2} \beta_{\tilde{B}}\Big(\bigwedge(\tilde{\mathbf{B}})\Big)_{\mathbf{B}=\mathbf{1},\mathbf{C}_2=\mathbf{0}} = \sum_{\tilde{\mathbf{B}}\subseteq\mathbf{B}} \beta_{\tilde{\mathbf{B}}}.$$

Namreč za množice $\tilde{\mathbf{B}}$, ki so vsebovane v $\mathbf{B}\cap\mathbf{C}_2$, velja enakost $(\bigwedge(\tilde{\mathbf{B}}))_{\mathbf{C}_2=\mathbf{0}} = 0$. Obenem veljajo tudi naslednje enakosti:

$$\begin{aligned} E[D \mid \mathbf{B} \setminus \{L\} = \mathbf{1},\ L = 0, \mathbf{C}_2 = \mathbf{0}] &= \sum_{\tilde{\mathbf{B}} \subseteq \mathbf{B} \cup \mathbf{C}_2} \beta_{\tilde{B}} \left(\bigwedge(\tilde{\mathbf{B}}) \right)_{\mathbf{B}\setminus\{L\}=\mathbf{1},\ L=0, \mathbf{C}_2=\mathbf{0}} \\ &= \sum_{\tilde{\mathbf{B}} \subseteq \mathbf{B}\setminus\{L\}} \beta_{\tilde{B}}. \end{aligned}$$

Nadaljujemo kot v dokazu izreka 7.1. □

Trditev 7.4 *Naj bo* $\boldsymbol{C} = \boldsymbol{C}_1 \dot\cup \boldsymbol{C}_2,\ \boldsymbol{B} = (\boldsymbol{B}_+ \dot\cup \boldsymbol{B}') \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}_1)),\ |\boldsymbol{B}| = |\boldsymbol{C}_1|$. *Naj* $\boldsymbol{B}_+ \cup \boldsymbol{C}_2$ *pozitivno monotono učinkuje na* D *glede na* $\boldsymbol{C}$. *Naslednja pogoja za neko drevo* $\mathfrak{T}$ *na* $\boldsymbol{B}_+$ *sta ekvivalentna:*

$$\begin{aligned} &(i) E[D \mid \boldsymbol{B} = \boldsymbol{1},\ \boldsymbol{C}_2 = \boldsymbol{0}] - \sum_{L \in \boldsymbol{B}} E[D \mid \boldsymbol{B} \setminus \{L\} = \boldsymbol{1},\ L = 0,\ \boldsymbol{C}_2 = \boldsymbol{0}] \\ &\quad + \sum_{\boldsymbol{E} \in \mathfrak{T}} E[D \mid \boldsymbol{B} \setminus \boldsymbol{E} = \boldsymbol{1},\ \boldsymbol{E} = \boldsymbol{0},\ \boldsymbol{C}_2 = \boldsymbol{0}] > 0. \\ &(ii) \sum_{\tilde{\boldsymbol{B}} \subseteq \boldsymbol{C}} m_{\tilde{\boldsymbol{B}}}^{ired} \beta_{\tilde{\boldsymbol{B}}} > 0,\ \textit{kjer} \\ &\quad m_{\tilde{\boldsymbol{B}}}^{ired} = 1 - |\boldsymbol{B} \setminus \tilde{\boldsymbol{B}}| + |\mathfrak{T}| - deg_{\mathfrak{T}}(\tilde{\mathbf{B}}) + |\mathfrak{T}_{\tilde{\mathbf{B}}}|. \end{aligned}$$

Dokaz: Veljajo naslednje enakosti:

$$\begin{aligned} E[D \mid \mathbf{B} \setminus \mathbf{E} = \mathbf{1},\ \mathbf{E} = \mathbf{0},\ \mathbf{C}_2 = \mathbf{0}] &= \sum_{\tilde{\mathbf{B}} \subseteq \mathbf{B} \cup \mathbf{C}_2} \beta_{\tilde{B}} \left(\bigwedge(\tilde{\mathbf{B}}) \right)_{\mathbf{B}\setminus\mathbf{E}=\mathbf{1},\ \mathbf{E}=\mathbf{0},\ \mathbf{C}_2=\mathbf{0}} \\ &= \sum_{\tilde{\mathbf{B}} \subseteq \mathbf{B}\setminus\mathbf{E}} \beta_{\tilde{B}}. \end{aligned}$$

Nadaljujemo podobno kot v dokazu trditve 7.2. □

Dokazali smo, da je pogoj $\sum_{\tilde{\mathbf{B}} \subseteq \mathbf{C}} m_{\tilde{\mathbf{B}}}^{ired} \beta_{\tilde{\mathbf{B}}} > 0$ zadosten pogoj minimalnosti.

7.3.4 Pogoji singularnosti pri predpostavki monotonosti

Poglejmo pogoj singularnosti pri predpostavki monotonosti, ki je ekvivalenten pogoju iz izreka 6.9.

Trditev 7.5 *Naj bo* $\boldsymbol{B} = \boldsymbol{C} = \boldsymbol{B}_+ \dot\cup \boldsymbol{B}' \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}))$ *in naj* $\boldsymbol{B}_+$ *pozitivno monotono učinkuje na* D *glede na* $\boldsymbol{C}$. *Naslednja pogoja za neko drevo* $\mathfrak{T}$ *na* $\boldsymbol{B}_+$ *sta ekvivalentna:*

$$\begin{aligned} &(i) E[D \mid \boldsymbol{B} = \boldsymbol{1}] - \sum_{L \in \boldsymbol{B}_+} E[D \mid \boldsymbol{B} \setminus \{L\} = \boldsymbol{1}, L = 0] \\ &- \sum_{\substack{\tilde{\boldsymbol{B}}:\ \tilde{\boldsymbol{B}} \subseteq \boldsymbol{B}' \\ \tilde{\boldsymbol{B}} \neq \emptyset}} E[D \mid \boldsymbol{B} \setminus \tilde{\boldsymbol{B}} = \boldsymbol{1}, \tilde{\boldsymbol{B}} = \boldsymbol{0}] + \sum_{\boldsymbol{E} \in \mathfrak{T}} E[D \mid \boldsymbol{B} \setminus \boldsymbol{E} = \boldsymbol{1}, \boldsymbol{E} = \boldsymbol{0}] > 0, \\ &(ii) \sum_{\tilde{\boldsymbol{B}} \subseteq \boldsymbol{C}} m_{\tilde{\boldsymbol{B}}}^{sing} \beta_{\tilde{\boldsymbol{B}}} > 0,\ \textit{kjer} \\ &\quad m_{\tilde{\boldsymbol{B}}}^{sing} \equiv m_{\tilde{\boldsymbol{B}}}^{ired} + |\boldsymbol{B}' \setminus \tilde{\boldsymbol{B}}| - (2^{|\boldsymbol{B}' \setminus \tilde{\boldsymbol{B}}|} - 1). \end{aligned}$$

Dokaz: Točneje velja $m_{\tilde{\mathbf{B}}}^{sing} \equiv 1 - |\mathbf{B} \setminus \tilde{\mathbf{B}}| + |\mathfrak{T}| - deg_{\mathfrak{T}}(\tilde{\mathbf{B}}) + |\mathfrak{T}_{\tilde{\mathbf{B}}}| + |\mathbf{B}' \setminus \tilde{\mathbf{B}}| - (2^{|\mathbf{B}' \setminus \tilde{\mathbf{B}}|} - 1)$. Najprej pokažimo, da velja naslednja enakost:

$$\sum_{L \in \mathbf{B}_+} E[D \mid \mathbf{B} \setminus \{L\} = \mathbf{1}, L = 0] = \sum_{\tilde{\mathbf{B}} \subseteq \mathbf{C}} (|\mathbf{B} \setminus \tilde{\mathbf{B}}| - |\mathbf{B}' \setminus \tilde{\mathbf{B}}|)\beta_{\tilde{\mathbf{B}}}.$$

Velja namreč

$$\begin{aligned}\sum_{L \in \mathbf{B}_+} E[D \mid \mathbf{B} \setminus \{L\} = \mathbf{1}, L = 0] &= \sum_{L \in \mathbf{B}_+} \sum_{\tilde{\mathbf{B}} \subseteq \mathbf{B} \setminus \{L\}} \beta_{\tilde{\mathbf{B}}} \\ &= \sum_{\tilde{\mathbf{B}} \subseteq \mathbf{B}} (|\mathbf{B}_+ \setminus \tilde{\mathbf{B}}|)\beta_{\tilde{\mathbf{B}}} \\ &= \sum_{\tilde{\mathbf{B}} \subseteq \mathbf{B}} (|\mathbf{B} \setminus \tilde{\mathbf{B}}| - |\mathbf{B}' \setminus \tilde{\mathbf{B}}|)\beta_{\tilde{\mathbf{B}}},\end{aligned}$$

kjer druga enakost sledi iz naslednjega razmišljanja: če $\tilde{\mathbf{B}} \subset \mathbf{B}'$ se koeficient $\beta_{\tilde{\mathbf{B}}}$ pojavi $|\mathbf{B}_+|$-krat; če $\tilde{\mathbf{B}} \subset \mathbf{B}^+$ se koeficient $\beta_{\tilde{\mathbf{B}}}$ pojavi $|\mathbf{B}_+ \setminus \tilde{\mathbf{B}}|$-krat; če pa $|\tilde{\mathbf{B}} \cap \mathbf{B}^+| \neq \emptyset$ se koeficient $\beta_{\tilde{\mathbf{B}}}$ pojavi $|\mathbf{B}_+ \setminus (\mathbf{B}^+ \cap \mathbf{B}^+)|$-krat. Torej je pri dani množici $\tilde{\mathbf{B}}$ število ponavljanj koeficienta $\beta_{\tilde{\mathbf{B}}}$ enako $|\mathbf{B}_+ \setminus (\mathbf{B}^+ \cap \mathbf{B}^+)| = |\mathbf{B}_+ \setminus \tilde{\mathbf{B}}|$.

Pokažimo še, da velja naslednja enačba:

$$\sum_{\substack{\tilde{\mathbf{B}}:\ \tilde{\mathbf{B}} \subseteq \mathbf{B}' \\ \tilde{\mathbf{B}} \neq \emptyset}} E[D \mid \mathbf{B} \setminus \tilde{\mathbf{B}} = \mathbf{1}, \tilde{\mathbf{B}} = \mathbf{0}] = \sum_{\tilde{\mathbf{B}} \subseteq \mathbf{C}} (2^{|\mathbf{B}' \setminus \tilde{\mathbf{B}}|} - 1)\beta_{\tilde{\mathbf{B}}}.$$

Velja namreč

$$\begin{aligned}\sum_{\substack{\tilde{\mathbf{B}}:\ \tilde{\mathbf{B}} \subseteq \mathbf{B}' \\ \tilde{\mathbf{B}} \neq \emptyset}} E[D \mid \mathbf{B} \setminus \tilde{\mathbf{B}} = \mathbf{1}, \tilde{\mathbf{B}} = \mathbf{0}] &= \sum_{\substack{\tilde{\mathbf{B}}:\ \tilde{\mathbf{B}} \subseteq \mathbf{B}' \\ \tilde{\mathbf{B}} \neq \emptyset}} \sum_{\mathbf{A} \subseteq \mathbf{B} \setminus \tilde{\mathbf{B}}} \beta_{\mathbf{A}} \\ &= \sum_{\mathbf{A} \subseteq \mathbf{B}} (2^{|\mathbf{B}' \setminus \mathbf{A}|} - 1)\beta_{\mathbf{A}},\end{aligned}$$

kjer druga enakost sledi iz dejstva, da je pri danem $\mathbf{A}$ je število pojavitev koeficienta $\beta_{\mathbf{A}}$ enako številu pravih podmnožic množice $\mathbf{B}' \setminus \mathbf{A}$. □

V naslednji posledici izrazimo pogoj singularnosti brez predpostavke monotonosti, ki sledi iz zgornje trditve.

Posledica 7.6 *Naj bo $\boldsymbol{B} \in \dot{\mathbb{P}}(\mathbb{L}(\boldsymbol{C}))$. Naslednja pogoja sta ekvivalentna:*

$$(i) E[D \mid \boldsymbol{B} = \boldsymbol{1}] - \sum_{\substack{\tilde{\boldsymbol{B}}:\ \tilde{\boldsymbol{B}} \subseteq \boldsymbol{B} \\ \tilde{\boldsymbol{B}} \neq \emptyset}} E[D \mid \boldsymbol{B} \setminus \tilde{\boldsymbol{B}} = \boldsymbol{1}, \tilde{\boldsymbol{B}} = \boldsymbol{0}] > 0,$$

$$(ii) \sum_{\tilde{\boldsymbol{B}} \subseteq \boldsymbol{C}} (2 - 2^{|\boldsymbol{B} \setminus \tilde{\boldsymbol{B}}|})\beta_{\tilde{\boldsymbol{B}}} > 0.$$

Dokaz: Sledi iz trditve 7.5 pri $\mathfrak{T} = \emptyset$ in $\mathbf{B} = \mathbf{B}'$. □

7.4 Preprostejši modeli

Podrobneje bomo obravnavali dva primera nasičenih Bernoullijevih regresijskih modelov. V prvem primeru bo model vključeval dve pojasnjevalni spremenljivki, v drugem pa tri pojasnjevalne spremenljivke. Ogledali si bomo teste iz prejšnjih poglavij v teh preprostejših modelih.

Primer 7.3 *Poglejmo nasičeno linearno Bernoullijevo regresijo z dvema pojasnjevalnima spremenljivkama X_1, X_2 iz primera 7.2. Iz trditve 7.1 sledi zadosten pogoj ireducibilnosti množice $\{X_1, X_2\}$:*

$$(1 - |\{X_1, X_2\}|)\beta_0 + (1 - |\{X_2\}|)\beta_1 + (1 - |\{X_1\}|)\beta_2 + (1 - |\emptyset|)\beta_{12} > 0,$$

oziroma

$$\beta_{12} > \beta_0.$$

Iz trditve 7.2 sledi pogoj ireducibilnosti množice $\{X_1, X_2\}$ pri pogoju monotonosti:

$$m_{\emptyset}^{ired}\beta_0 + m_{X_1}^{ired}\beta_1 + m_{X_2}^{ired}\beta_2 + m_{X_1,X_2}^{ired}\beta_{12} > 0,$$

oziroma

$$\begin{aligned}
&(1 - |\{X_1, X_2\} \setminus \emptyset| + |\mathfrak{T}| - deg_{\mathfrak{T}}(\emptyset) + |\mathfrak{T}_{\emptyset}|)\beta_0 \\
&+ (1 - |\{X_1, X_2\} \setminus \{X_1\}| + |\mathfrak{T}| - deg_{\mathfrak{T}}(\{X_1\}) + |\mathfrak{T}_{\{X_1\}}|)\beta_1 \\
&+ (1 - |\{X_1, X_2\} \setminus \{X_2\}| + |\mathfrak{T}| - deg_{\mathfrak{T}}(\{X_2\}) + |\mathfrak{T}_{\{X_2\}}|)\beta_2 \\
&+ (1 - |\{X_1, X_2\} \setminus \{X_1, X_2\}| + |\mathfrak{T}| - deg_{\mathfrak{T}}(\{X_1, X_2\}) + |\mathfrak{T}_{\{X_1,X_2\}}|)\beta_{12} > 0.
\end{aligned}$$

V primeru, da le X_1 pozitivno monotono učinkuje na D, potem je drevo $\mathfrak{T}$ enako prazni množici, $\mathfrak{T} = \emptyset$, zato je zadosten pogoj ireducibilnosti množice $\{X_1, X_2\}$ enak zgornjemu pogoju $\beta_{12} > \beta_0$. Ta pogoj je po trditvi 6.7 tudi zadosten pogoj singularnosti pri monotonem učinku X_1. Če oba X_1 in X_2 pozitivno monotono učinkujeta na D, potem je drevo enako $\mathfrak{T} = \{X_1, X_2\}$ in je pogoj ireducibilnosti enak

$$\beta_{12} > 0.$$

Zadnja neenakost je po trditvi 6.7 tudi zadosten pogoj singularnosti pri monotonem učinku X_1, X_2.

Iz trditve 4.4 sledi, da je pogoj minimalnosti množice $\{X_1, X_2\}$ enak pogoju ireducibilnosti.

V primeru, da ni nobenega pozitivnega učinka, je po posledici 7.6 zadosten pogoj singularnosti množice $\{X_1, X_2\}$ enak pogoju

$$\begin{aligned}
&(2 - 2^{|\{X_1,X_2\}\setminus\emptyset|})\beta_0 + (2 - 2^{|\{X_1,X_2\}\setminus\{X_1\}|})\beta_1 \\
&+ (2 -^{|\{X_1,X_2\}\setminus\{X_2\}|})\beta_2 + (2 - 2^{|\{X_1,X_2\}\setminus\{X_1,X_2\}|})\beta_{12} > 0,
\end{aligned}$$

oziroma

$$\beta_{12} > 2\beta_0.$$

Primer 7.4 *Obravnavajmo še nasičen linearen Bernoullijev regresijski model s tremi binarnimi pojasnjevalnimi spremenljivkami $X_1,\ X_2,\ X_3$:*

$$\begin{aligned}
E[D \mid X_1 = x_1, X_2 = x_2, X_3 = x_3] =& \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 \\
&+ \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3.
\end{aligned}$$

Predpostavimo, da množice $\{X_1, X_2, X_3\}$ ne učinkujejo zavajajoče na D. Iz trditve 7.1 sledi zadosten pogoj ireducibilnosti množice $\{X_1, X_2, X_3\}$ brez predpostavke monotonosti:

$$\beta_{123} > 2\beta_0 + \beta_1 + \beta_2 + \beta_3. \tag{7.4}$$

Poglejmo pogoje pri predpostavki monotonosti. Iz trditve 7.2 sledi pogoj ireducibilnosti pri predpostavki monotonosti:

$$\begin{aligned} m_{\emptyset}^{ired}\beta_0 + m_{X_1}^{ired}\beta_1 + m_{X_2}^{ired}\beta_2 + m_{X_3}^{ired}\beta_3 + m_{X_1,X_2}^{ired}\beta_{12} \\ + m_{X_1X_3}^{ired}\beta_{13} + m_{X_2X_3}^{ired}\beta_{23} + m_{X_1X_2X_3}^{ired}\beta_{123} > 0, \end{aligned}$$

kjer pri monotonem učinku X_1, X_2 in X_3 velja

$$\begin{aligned} m_{\tilde{\boldsymbol{B}}}^{ired} &= 1 - |\boldsymbol{B} \setminus \tilde{\boldsymbol{B}}| + |\mathfrak{T}| - deg_{\mathfrak{T}}(\tilde{\mathbf{B}}) + |\mathfrak{T}_{\tilde{\mathbf{B}}}| \\ &= \begin{cases} 1-3+2-0+0 & \textit{če } |\tilde{\boldsymbol{B}}| = \emptyset \\ 1-2+2-1+0 & \textit{če } |\tilde{\boldsymbol{B}}| = 1,\ deg_{\mathfrak{T}}(\tilde{\mathbf{B}}) = 1 \\ 1-2-2-2+0 & \textit{če } |\tilde{\boldsymbol{B}}| = 1,\ deg_{\mathfrak{T}}(\tilde{\mathbf{B}}) = 2 \\ 1-1+2-3+1 & \textit{če } |\tilde{\boldsymbol{B}}| = 2 \\ 1-0+2-4+2 & \textit{če } |\tilde{\boldsymbol{B}}| = 3 \end{cases} \\ &= \begin{cases} 0 & \textit{če } |\tilde{\boldsymbol{B}}| = \emptyset \\ 0 & \textit{če } |\tilde{\boldsymbol{B}}| = 1,\ deg_{\mathfrak{T}}(\tilde{\mathbf{B}}) = 1 \\ -1 & \textit{če } |\tilde{\boldsymbol{B}}| = 1,\ deg_{\mathfrak{T}}(\tilde{\mathbf{B}}) = 2 \\ 0 & \textit{če } |\tilde{\boldsymbol{B}}| = 2 \\ 1 & \textit{če } |\tilde{\boldsymbol{B}}| = 3. \end{cases} \end{aligned}$$

Sledijo trije zadostni pogoji ireducibilnosti množice $\{X_1, X_2, X_3\}$ pri monotonem učinku X_1, X_2 in X_3:

$$\beta_{123} > \beta_1 \quad \textit{če izberemo} \quad \mathfrak{T} = \{\{X_1, X_2\}, \{X_1, X_3\}\}, \tag{7.5a}$$

$$\beta_{123} > \beta_2 \quad \textit{če izberemo} \quad \mathfrak{T} = \{\{X_1, X_2\}, \{X_2, X_3\}\}, \tag{7.5b}$$

$$\beta_{123} > \beta_3 \quad \textit{če izberemo} \quad \mathfrak{T} = \{\{X_1, X_3\}, \{X_2, X_3\}\}. \tag{7.5c}$$

Če pa le X_1 in X_2 pozitivno monotono učinkujeta na D, potem velja

$$\begin{aligned} m_{\tilde{\boldsymbol{B}}}^{ired} &= 1 - |\boldsymbol{B} \setminus \tilde{\boldsymbol{B}}| + |\mathfrak{T}| - deg_{\mathfrak{T}}(\tilde{\mathbf{B}}) + |\mathfrak{T}_{\tilde{\mathbf{B}}}| \\ &= \begin{cases} 1-3+1-0+0 & \textit{če } |\tilde{\boldsymbol{B}}| = \emptyset \\ 1-2+1-0+0 & \textit{če } |\tilde{\boldsymbol{B}}| = 1,\ deg_{\mathfrak{T}}(\tilde{\mathbf{B}}) = 0 \\ 1-2+1-1+0 & \textit{če } |\tilde{\boldsymbol{B}}| = 1,\ deg_{\mathfrak{T}}(\tilde{\mathbf{B}}) = 1 \\ 1-1+1-2+1 & \textit{če } |\tilde{\boldsymbol{B}}| = 2,\ \tilde{\boldsymbol{B}} = \{X_1, X_2\} \\ 1-1+1-1+0 & \textit{če } |\tilde{\boldsymbol{B}}| = 2,\ \tilde{\boldsymbol{B}} \neq \{X_1, X_2\} \\ 1-0+1-2+1 & \textit{če } |\tilde{\boldsymbol{B}}| = 3 \end{cases} \end{aligned} \tag{7.6}$$

$$= \begin{cases} -1 & \textit{če } |\tilde{\boldsymbol{B}}| = \emptyset \\ 0 & \textit{če } |\tilde{\boldsymbol{B}}| = 1,\ deg_{\mathfrak{T}}(\tilde{\mathbf{B}}) = 0 \\ -1 & \textit{če } |\tilde{\boldsymbol{B}}| = 1,\ deg_{\mathfrak{T}}(\tilde{\mathbf{B}}) = 1 \\ 0 & \textit{če } |\tilde{\boldsymbol{B}}| = 2,\ \tilde{\boldsymbol{B}} = \{X_1, X_2\} \\ 0 & \textit{če } |\tilde{\boldsymbol{B}}| = 2,\ \tilde{\boldsymbol{B}} \neq \{X_1, X_2\} \\ 1 & \textit{če } |\tilde{\boldsymbol{B}}| = 3. \end{cases}$$

Sledi zadosten pogoj ireducibilnosti množice $\{X_1, X_2, X_3\}$ pri pozitivnem učinku X_1 in X_2 :

$$\beta_{123} > \beta_0 + \beta_1 + \beta_2. \tag{7.7}$$

Kot pričakovano je najstrožji pogoj ireducibilnosti pogoj (7.4) brez predpostavke monotonosti, šibkejši od njega je pogoj (7.7) pri predpostavki pozitivnega učinka X_1 in X_2, najšibkejši pa je katerikoli izmed pogojev (7.5) pri predpostavki pozitivnega učinka X_1, X_2 in X_3. Monoton učinek nam namreč omogoča, da omilimo pogoje.

V primeru, da le en dogodek pozitivno monotono učinkuje na D, lahko uporabimo le pogoj (7.4) pri preverjanju ireducibilnosti, saj ne moremo zgraditi drevesa na enem samem dogodku.

V zgornjih primerih, kjer vsi ali vsi razen enega dejavnika učinkujejo monotono, je pogoj ireducibilnosti enak pogoju singularnosti. Zadnje sledi iz trditve 6.7.

Iz posledice 7.6 sledi zadosten pogoj singularnosti množice $\{X_1, X_2, X_3\}$ brez predpostavke monotonosti:

$$\begin{aligned}&(2-2^3)\beta_0 + (2-2^2)\beta_1 + (2-2^2)\beta_2 + (2-2^2)\beta_3\\&+ (2-2^1)\beta_{12} + (2-2^1)\beta_{23} + (2-2^1)\beta_{13} + (2-2^0)\beta_{123} > 0,\end{aligned}$$

oziroma

$$\beta_{123} > 6\beta_0 + 2\beta_1 + 2\beta_2 + 2\beta_3.$$

Iz trditve 4.4 sledi še, da so pogoji minimalnosti enaki pogojem ireducibilnosti.

8 Interakcija genotipov s kajenjem

Najbolj znani vzroki raka na sečnem mehurju so izpostavljenosti okoljskim razmeram, predvsem pa izpostavljenost tobačnemu dimu in kajenje. Raziskave so pokazale, da sta pomembna dejavnika pri raziskovanju raka na sečnem mehuru tudi encima imenovana *N-acetyltransferase 1* (NAT1) in *N-acetyltransferase 2* (NAT2).

Namreč, pomembna sestavina tobačnega dima je tudi rakotvorna snov *arilamine*. Encima NAT1 in NAT2 pa imata pomembno vlogo pri metaboliziranju rakotvornih snovi v sečnem mehurju.

Encim NAT1 je sestavljen iz različnih skupin genov. Te skupine imenujemo *genotipi*. Genotipi encima NAT1 se razlikujejo po vsebini različnih delcev, ki jih imenujemo *aleli*. Ti so na primer NAT*3, NAT*4, NAT*11 in NAT*10. Aleli v NAT2 pa so na primer NAT2*4, NAT*5, NAT*6, NAT*7 in NAT*14.

Genotip encima NAT2, ki vsebuje enega ali več kopij alela NAT*4, imenujemo R NAT2; genotip encima NAT2, ki pa vsebuje katerokoli kombinacijo alelov NAT*5, NAT*6, NAT*7 in NAT*14 ter ne vsebuje alela NAT*4, imenujemo S NAT2.

V raziskavi [9] o interakcijah med posameznimi genotipi NAT1 in NAT2 ter kajenjem so ugotovili, da prisotnost alela NAT1*10, alela iz skupine S NAT2 ter kajenja poveča nevarnost raka na sečnem mehurju. Interakcija zgornjih treh dejavnikov je po rezultatih raziskave statistično pomembna. Analizirali bomo interakcije dejavnikov tudi s pomočjo obravnavane teorije.

Najprej poglejmo, kako bomo označevali zgornje dejavnike. Naj spremenljivka izidov D označuje rak sečnega mehurja: $D(\omega) = 1$, če ima osebek ω raka na sečnem mehurju, ter $D(\omega) = 0$, če osebek nima raka na sečnem mehurju. Naj spremenljivka X_1 označuje kajenje: $X_1(\omega) = 1$, če je osebek ω kadilec, ter $X_1(\omega) = 0$, če je nekadilec. Spremenljivka X_2 označuje genetično različico genotipa NAT2: $X_2(\omega) = 1$, če ima osebek ω genotip S NAT2, ter $X_2(\omega) = 0$, če pa ima osebek genotip R NAT2. Naj spremenljivka X_3 označuje prisotnost alela NAT1*10 v genotipu encima NAT1: $X_3(\omega) = 1$, če ima osebek ω alel NAT1*10 v genotipu encima NAT1, ter $X_3(\omega) = 0$, če osebek nima alela NAT1*10 v genotipu encima NAT1. Zgornje spremenljivke definiramo na populaciji Ω sestavljeni iz osebkov kavkazijske rase.

Obravnavajmo nasičen regresijski model kot v primeru 7.4:

$$E[D \mid X_1 = x_1, X_2 = x_2, X_3 = x_3] = \beta_\emptyset + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3.$$

Predpostavimo, da ni zavajajočega učinka $\mathbf{C} = \{X_1, X_2, X_3\}$ na D ter da imajo X_1, X_2 in X_3 pozitiven monoton učinek na D. V primeru 7.4 smo ugotovili, da je vsak izmed naslednjih pogojev zadosten pogoj ireducibilnosti množice $\{X_1, X_2, X_3\}$:

$$\beta_{123} > \beta_1, \qquad \beta_{123} > \beta_2, \qquad \beta_{123} > \beta_3.$$

V članku [10] avtorja izračunata zgornje regresijske koeficiente s pomočjo *razmerja obetov* (odds ratio), ki ga definiramo na sledeči način:

$$OR_{x_1,x_2,x_3} = O_{x_1,x_2,x_3} / O_{0,0,0},$$

kjer je *obet* izida

$$O_{x_1,x_2,x_3} = \frac{P[D = 1 \mid X_1 = x_1, X_2 = x_2, X_3 = x_3]}{1 - P[D = 1 \mid X_1 = x_1, X_2 = x_2, X_3 = x_3]}.$$

V sledeči tabeli si lahko ogledamo vrednosti pojasnjevalnih spremenljivk in ocene razmerij obetov v posameznih podskupinah.

Tabela 2: *Podatki iz študije raka sečnega mehurja.*

Kadilec	**S NAT2**	**NAT1*10**	**Zdravljena skupina (n = 215)**	**Kontrolna skupina (n = 191)**	**Razmerje obetov (95% CI)**
0	0	0	6	13	1
0	0	1	8	16	1.1 (0.3, 3.9)
0	1	0	16	31	1.1 (0.4, 3.5)
0	1	1	6	10	1.3 (0.3, 5.3)
1	0	0	42	32	2.8 (1.1, 8.3)
1	0	1	41	26	3.4 (1.2, 10.1)
1	1	0	61	51	2.6 (0.9, 7.3)
1	1	1	35	12	6.3 (2.0, 20.3)

Avtorja članka [10] prideta do naslednjih zaključkov: če ni zavajajočega učinka $\mathbf{C}$ na D ter imajo dogodki X_1, X_2 in X_3 pozitiven monoton učinek na D, potem rezultati namigujejo na ireducibilnost množice X_1, X_2, X_3, vendar ne moremo zagovarjati ireducibilnosti množice, saj velikost vzorca ni dovolj velika. Podoben zaključek dobimo v primeru, ko imata le dva dogodka izmed X_1, X_2 in X_3 pozitiven monoton učinek na D. Če ne predpostavimo, da je prisoten katerikoli monoton učinek, potem rezultati sploh ne namigujejo na ireducibilnost množice.

Literatura

[1] J.D. Angrist, J.S. Pischke, *Mostly harmless econometrics: An empiricisti's companion*, Princeton University Press, Princeton, New Jersey (2008).

[2] A. Cayley, *A theorem on trees*. Quart. J. Math. **23** (1889) 376-378.

[3] R. Dudley, *The delta-method, multinomial distributions, and en example: standard error of log odds ratios*, Statistics for Applications 18.443 (2009) (Massachusetts Institute of Technology: MIT OpenCourseWare), [ogled 2. 8. 2013], dostopno na `http://ocw.mit.edu/courses/mathematics/18-443-statistics-for-applications-spring-2009/assignments/MIT18_443s09_assn07_res01.pdf`. License: Creative Commons BY-NC-SA.

[4] R.H. Hoyle, *Structural equation modeling: concepts, issues and applications*, SAGE Publications, California (1995).

[5] A.R. Jiang, C.M. Gao, J.H. Ding, S.P. Li, Y.T. Liu, H.X. Cao, J.Z. Wu, J.H. Tang, Y. Quan, K. Tajima, *Abortion and breast risk cancer risk in promenopausal and postmenopausal women in Jangsu province of China*, Asian Pacific J Cancer Prev **13** (2012) 33-35.

[6] W.W. LaMorte, *The sufficient-component cause model*, Boston University School of Public Health (2012), [ogled 13. 6. 2013], dostopno na `http://sph.bu.edu/otlt/MPH-Modules/EP/EP713_Causality/EP713_Causality4.html`.

[7] S.L. Morgan, C. Winship, *Counterfactuals and causal inference, methods and principles for social research*, Cambridge University Press, New York (2007).

[8] J.A. Rice, *Mathematical statistics and data analysis*, 3rd edition, Duxbury Advanced Series, Belmont, (2007).

[9] J.A. Taylor, D.M. Umbach, E. Stephens, T. Castranio, D. Paulson, C. Robertson, JL. Mohler, D.A. Bell, *The role of N-acetylation polymorphisms in smoking-associated bladder cancer: evidence of a gene-gene-exposure three-way interaction*, Cancer Research **58** (1998) 53603–3610.

[10] T.J. Vanderweele, T.S. Richardson, *General theory for interactions in sufficient cause models with dichotomous exposures*, The Annals od Statistics **40**(4) (2012) 2128-2161.

[11] T.J. Vanderweele, J.M. Robins, *Empirical and counterfactual conditions for sufficient cause interactions*, Biometrika **95**(1) (2008) 49-61.