

15

**Richard Scheines*****Introduzione all'inferenza causale***

Scheines illustra le caratteristiche dei grafi aciclici orientati, presentandone alcuni aspetti fondamentali quali la "d-separazione", la "condizione markoviana causale", la "condizione di fedeltà". Dopo avere insistito sulla distinzione tra i grafi in quanto oggetti matematici e la loro interpretazione causale, Scheines conclude sottolineando come, contrariamente alla lettura che alcuni hanno dato dell'approccio elaborato da lui, Spirtes e Glymour, la teoria non ha affatto la pretesa di «far uscire conigli causali da un cappello statistico» [Scheines 1997, 185-199].

*In Causation, Prediction and Search* Peter Spirtes, Clark Glymour e io abbiamo sviluppato una teoria dell'inferenza causale statistica. Cercherò di presentare quali siano le nostre intuizioni, sottolineando alcune delle idee fondamentali e fornendo esempi particolarmente semplici. Il percorso che sceglierò di seguire vuole essere la risposta alle perplessità di molti. Le nostre tecniche operano su dati statistici e producono serie di grafi orientati. La maggior parte degli studiosi interessati vedono come queste tecniche funzionano, ma non comprendono facilmente le assunzioni aggiuntive necessarie per dare a questi risultati un'interpretazione causale, ovvero un'interpretazione che ci dica come i sistemi risponderebbero a degli interventi. Cercherò di presentare nel modo più semplice possibile le assunzioni che ci permettono di passare da relazioni di indipendenza probabilistica al tipo di relazioni causali che comportano controfattuali relativi a manipolazioni e interventi. Separerò anzitutto le varie parti della teoria: i grafi orientati, la probabilità e la causalità, e chiarirò poi le assunzioni che legano la struttura causale alla probabilità. Discuterò, infine, le assunzioni aggiuntive necessarie per compiere inferenze dai dati statistici alla struttura causale.

***1. Grafi aciclici orientati e d-separazione***

La teoria che abbiamo elaborato coniuga due parti di matematica e una parte di filosofia. Le parti di matematica sono i grafi aciclici orientati (DAG) e la teoria della probabilità (con un interesse soprattutto per l'indipendenza condizionale); la parte di filosofia concerne le relazioni causali tra variabili.

Un DAG è un insieme di vertici e un insieme di archi (freccie) che uniscono coppie di questi vertici. Ad esempio, possiamo avere un insieme di tre vertici:  $\{X_1, X_2, X_3\}$ , e un insieme di due frecce che uniscono tali vertici:  $\{X_1 \rightarrow X_2; X_2 \rightarrow X_3\}$ . Rappresentiamo quasi sempre i DAG con un disegno o un diagramma. Questo DAG, ad esempio, può essere rappresentato così:

$$X_1 \rightarrow X_2 \rightarrow X_3$$

Prima di qualunque interpretazione, un DAG è un oggetto matematico completamente astratto. Nella nostra teoria, ai DAG sono attribuite due funzioni distinte. In rapporto alla prima rappresentano insiemi di distribuzioni di probabilità; in rapporto alla seconda rappresentano strutture causali. Il modo in cui rappresentano distribuzioni di probabilità è dato dalla condizione markoviana, che nei DAG risulta equivalente a una relazione grafica più utile in generale: la d-separazione [Pearl 1988]. La d-separazione è una relazione tra tre insiemi disgiunti di vertici in un grafico orientato. Sebbene sia troppo complicata per spiegarla o definirla qui, l'idea fondamentale è quella di controllare se un insieme di vertici  $Z$  blocca tutte le connessioni di un certo tipo tra  $X$  e  $Y$  in un grafo  $G$ . Se è così, allora  $X$  e  $Y$  sono d-separati da  $Z$  in  $G$ .

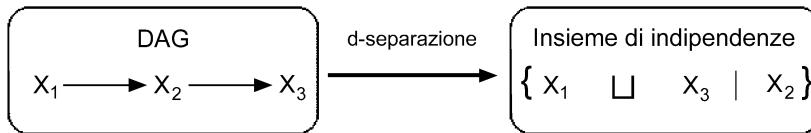


Fig. 1

Nel DAG a sinistra nella figura 1, ad esempio,  $X_2$  blocca l'unico percorso orientato che connette  $X_1$  e  $X_3$ , quindi  $X_1$  e  $X_3$  sono d-separati da  $X_2$  in questo DAG. Scegliendo la d-separazione per connettere i DAG a delle distribuzioni di probabilità, assumiamo che in tutte le distribuzioni  $P$  che un DAG  $G$  può rappresentare, se insiemi di vertici  $X$  e  $Y$  sono d-separati da un insieme  $Z$  nel DAG  $G$ , allora  $X$  e  $Y$  sono indipendenti condizionati a  $Z$  in  $P$ . Ad esempio, applicando la d-separazione al DAG nella figura 1 abbiamo:  $X_1$  e  $X_3$  sono d-separati da  $X_2$ . Assumiamo allora che in tutte le distribuzioni che questo DAG può rappresentare,  $X_1$  è indipendente da  $X_3$  condizionato a  $X_2$ . Per rappresentare l'indipendenza usiamo una notazione introdotta da Phil Dawid [1979];  $X_1 \perp X_3 \mid X_2$  significa:  $X_1$  e  $X_3$  sono indipendenti condizionati a  $X_2$ .

Si deve sottolineare che finché restiamo agnostici e non diamo alcuna interpretazione dei DAG, questi ultimi sono solo oggetti matematici che possiamo mettere in relazione a distribuzioni probabilistiche in qualunque modo vogliamo. Potremmo definire e usare altrettanto facilmente una e-separazione, o una f-separazione, o qualsiasi relazione grafica desideriamo, purché produca insiemi coerenti di indipendenze. Quando diamo un'interpretazione causale dei DAG, allora diviene necessario sostenere che la d-separazione è la connessione corretta tra un DAG causale e le distribuzioni di probabilità. Posticipiamo, comunque, questo tema di alcune pagine.

Spesso ci sono molti DAG distinti che rappresentano esattamente il medesimo insieme di relazioni di indipendenza, e dunque lo stesso insieme di distribuzioni. E così come è desiderabile avere una procedura che calcola la d-separazione per ogni grafo, è auspicabile avere un algoritmo che calcoli tutti i DAG che rappresentano un certo insieme di relazioni di indipendenza.

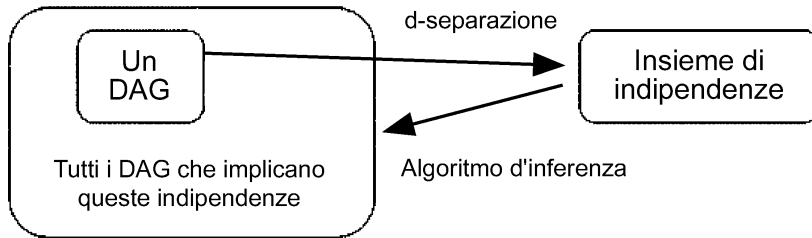


Fig. 2

Abbiamo sviluppato parecchi di tali algoritmi. Uno di questi è chiamato “algoritmo PC” ed è calcolato dal programma TETRAD II. Il suo input è un insieme di relazioni di indipendenza su un insieme di variabili, e il suo output è un insieme di DAG sulle variabili che sono equivalenti dal punto di vista della d-separazione, o della condizione markoviana. Applicando l'algoritmo PC allo stesso insieme di relazioni di indipendenza illustrato nel lato destro della figura 1, è possibile vedere (figura 3) che ci sono altri due DAG che sono equivalenti dal punto di vista della d-separazione al DAG della figura 1. PC è noto per essere completo, nel senso che il suo input include tutti e solo quei DAG che sono equivalenti dal punto di vista della d-separazione.

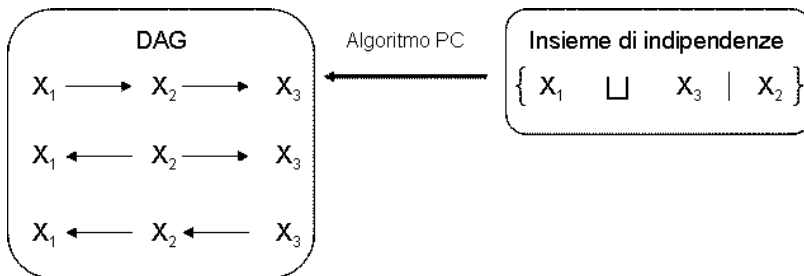


Fig. 3

## 2. Grafi causali

Se non si va oltre, la d-separazione e la condizione markoviana sono soltanto matematica che connette i DAG alle distribuzioni di probabilità, e non hanno affatto bisogno di coinvolgere la causalità. Ci si può accontentare di usare questa teoria matematica per produrre semplicemente rappresentazioni compatte ed eleganti di strutture di indipendenza, oppure si può compiere un passo ulteriore, assumendo che quando i DAG vengono *interpretati causalmente* la condizione markoviana e la d-separazione costituiscono la connessione corretta tra la struttura causale e l'indipendenza probabilistica. Chiamiamo quest'ultima assunzione condizione markoviana causale, e si tratta di un'assunzione più forte della condizione markoviana.

I DAG che vengono interpretati causalmente sono chiamati *grafi causali*. C'è una freccia da  $X$  a  $Y$  in un grafo causale che include un insieme di

variabili  $V$  solo nel caso in cui  $X$  sia una causa diretta di  $Y$  relativamente a  $V$ . Ad esempio, se  $S$  è una variabile che sta per l'abitudine a fumare,  $Y$  una variabile che sta per dita ingiallite, o macchiate dalla nicotina, e  $C$  una variabile che sta per la presenza di un tumore ai polmoni, allora il grafo causale nella figura 4 rappresenta quella che credo essere la struttura causale sussistente tra queste variabili.

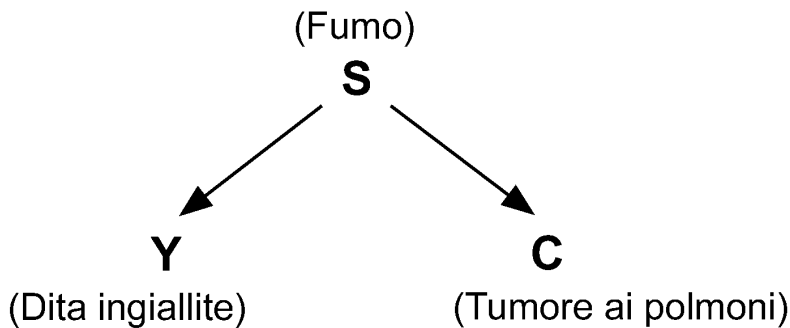


Fig. 4

Si assume che i grafi causali siano completi in un certo senso, e non in un altro. Sono incompleti nel senso che non includono necessariamente *tutte* le cause di ciascuna variabile del sistema. Pertanto molte delle cause del tumore sono state lasciate fuori dal grafo, ad esempio, l'inalazione di amianto, fattori genetici, ecc. Sono lasciate fuori anche molte variabili che possono trovarsi tra una certa causa specifica e il suo effetto; ad esempio un trauma nel rivestimento bronchiale può situarsi lungo il percorso causale "vero" che va dal fumo al tumore ai polmoni. Ma si assume che un grafo causale sia completo nel senso che tutte le *cause comuni* delle variabili specificate sono incluse. Ad esempio, se c'è una qualche variabile che è una causa tanto dell'abitudine a fumare quanto del tumore ai polmoni, ad esempio un certo fattore genetico, allora il grafo causale riportato nella figura 4 non è una rappresentazione accurata della struttura causale di queste tre variabili. Si assume che il grafo causale sia completo anche nel senso che *tutte* le relazioni causali tra le variabili specificate sono incluse nel grafo. Ad esempio, il grafo nella figura 4 non include una freccia che vada da  $Y$  a  $S$ , pertanto è accurato solo se le macchie di nicotina non causano in alcun modo l'abitudine a fumare.

La semantica di un grafo causale include delle manipolazioni ideali e i cambiamenti nella distribuzione di probabilità che seguono tali manipolazioni. Un resoconto di questo tipo è circolare, giacché manipolare è causare. Il nostro scopo, in ogni caso, non è fornire una definizione riduzionista della causalità, bensì, piuttosto, connetterla alla probabilità in un modo che rispecchi la pratica scientifica e permetta un'indagine sistematica sull'inferenza causale.

Manipolare una variabile è, idealmente, mutarla in un modo che, almeno

per il momento, lasci indisturbate tutte le altre variabili. Una manipolazione di questo genere deve cambiare direttamente solo la variabile sulla quale viene esercitata, e lasciare che i cambiamenti nelle altre variabili vengano prodotti da questo primo cambiamento, o che non vengano prodotti affatto. Ad esempio, supponiamo di cercare di controllare sperimentalmente delle ipotesi concernenti le relazioni causali che sussistono tra un'esibizione atletica e la sicurezza in sé che nutre l'atleta. Supponiamo di intervenire per minare la performance atletica somministrando un farmaco che blocca l'assorbimento delle sostanze nutrienti da parte delle cellule muscolari. Questa droga, però, imita anche la struttura chimica dei neurotrasmettitori che inibiscono il senso di insicurezza e ansia, e serve dunque direttamente ad aumentare l'ansia e a diminuire la fiducia che l'atleta nutre nelle sue capacità. Questo intervento aiuta ben poco a riflettere sul tipo di relazione causale che sussiste tra la performance atletica e l'autostima, poiché altera direttamente entrambe le variabili. È un intervento "con la mano pesante". Gli interventi ideali sono perfettamente selettivi rispetto alle variabili che cambiano *direttamente*.

Qualunque manipolazione ideale possiamo considerare, i grafi causali ci dicono quali variabili ci aspettiamo di mutare in qualche modo, e quali no. In parole semplici, le sole variabili che possiamo sperare di cambiare devono essere situate "a valle" delle variabili che manipoliamo. Anche se possiamo fare delle inferenze "a monte", ovvero dagli effetti alle cause, non possiamo manipolare un effetto e sperare di cambiare così le sue cause. Nella figura 4, ad esempio, dopo una manipolazione ideale delle macchie di nicotina, nulla accadrà alle probabilità di fumare e di essere colpiti da un tumore ai polmoni. Queste ultime avranno lo stesso valore che avrebbero avuto se non avessimo compiuto alcuna manipolazione. Se potessimo manipolare la probabilità di un certo individuo di essere colpito da un tumore ai polmoni senza perturbare direttamente la sua abitudine a fumare o le sue dita ingiallite, allora, di nuovo, non aspetteremmo di mutare la probabilità che l'individuo ha di fumare o di avere dita macchiate.

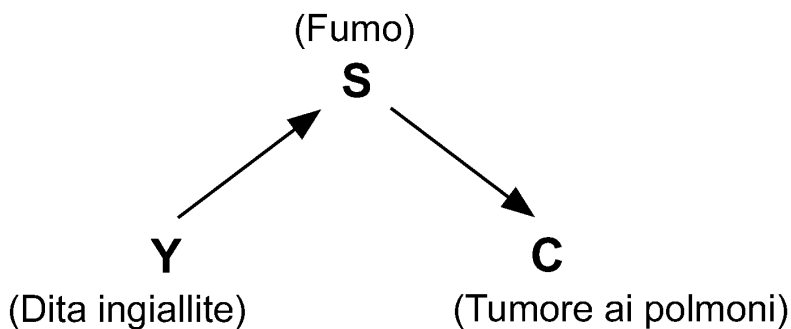


Fig. 5

Se il grafo causale cambia, lo stesso fa l'insieme di controfattuali concernenti le manipolazioni ideali. Se, ad esempio, abbiamo un grafo causale come quello della figura 5 (per quanto possa sembrare assurdo), allora solo la proposizione concernente la manipolazione del tumore ai polmoni resta immutata. Nessuna manipolazione ideale del fumo sfocerà in un cambiamento di  $Y$ , mentre alcune daranno come esito un cambiamento della probabilità di  $C$ , e alcune manipolazioni di  $Y$  avranno come esito dei cambiamenti della probabilità di  $S$ .

Senza addentrarci in tutte le sottigliezze del caso, il punto fondamentale dovrebbe essere chiaro: il tipo di causalità sul quale indaghiamo risponde a un sistema di interventi.

### 3. La condizione markoviana causale

L'assunzione markoviana causale può essere espressa in modo semplice.

Una variabile  $X$  è indipendente da ogni altra variabile (ad eccezione degli effetti di  $X$ ) condizionata a tutte le sue cause dirette.

Applicandola a ciascuna variabile nel grafo causale nella figura 4 otteniamo le seguenti relazioni di indipendenza:

Per  $Y$ : la variabile  $Y$  è indipendente da  $C$  condizionata a  $S$ .

Per  $S$ : Tutte le altre variabili sono effetti di  $S$ , dunque la condizione è vacua.

Per  $C$ : la variabile  $C$  è indipendente da  $Y$  condizionata a  $S$ .

Per la teoria della probabilità, la prima e l'ultima di queste indipendenze sono equivalenti, quindi questo grafo causale implica un'indipendenza secondo l'assunzione markoviana causale. Si può vedere che la figura 5 implica le stesse relazioni di indipendenza della 4, anche se il grafo è diverso dal punto di vista causale e pertanto implica diversi controfattuali concernenti gli interventi.

Le relazioni di indipendenza che risultano applicando l'assunzione markoviana causale a un grafo causale sono le stesse che si ottengono applicando la  $d$ -separazione a un grafo causale, ma è più semplice giustificare la connessione tra grafi causali e probabilità se la si espone in forma markoviana. L'intuizione sottesa all'assunzione markoviana causale è semplice: ignorando gli effetti di una variabile, tutta l'informazione probabilistica rilevante in merito a questa che può essere ottenuta da un sistema è contenuta nelle sue cause dirette. In un processo markoviano, conoscere lo stato presente di un sistema è rilevante per il suo futuro, mentre conoscere come è pervenuto al suo stato presente è del tutto irrilevante. Hans Reichenbach [1956] è stato il primo filosofo a discutere esplicitamente le proprietà markoviane dei sistemi causali, ma alcune varianti sono state discusse anche da Nancy Cartwright [1989], Wesley Salmon [1984], Brian Skyrms [1980], Patrick Suppes [1970], e molti altri.

Come fa un'assunzione di questo tipo a cogliere l'asimmetria della causalità? Per sistemi di due variabili non può farlo. I due grafi causali nella figura 6 per la condizione markoviana causale implicano le stesse indipendenze, e sulla base della sola indipendenza probabilistica risultano quindi indistinguibili. Ma saltare da questa semplice impossibilità di distinguerli alla conclusione che le probabilità non ci possono mai fornire informazione in merito alla struttura causale è chiaramente fallace. Come Hausman [1984] e Papineau [1985] hanno compreso, aggiungere una terza variabile cambia completamente la faccenda. I due grafi nella figura 7 *non* sono equivalenti dal punto di vista della relazione markoviana o della d-separazione, e la differenza tra le loro implicazioni di indipendenza è più in generale sottesa alla connessione tra indipendenza e priorità causale.



Fig. 6



Relazioni di indipendenza implicate dalla d-separazione

$$X \perp\!\!\!\perp Z \qquad X \perp\!\!\!\perp Z \mid Y$$

Fig. 7

Vediamo tre principali linee di giustificazione per l'assunzione markoviana causale, sebbene ce ne siano certamente anche altre. Primo, delle versioni dell'assunzione vengono adoperate, forse implicitamente, nel compiere inferenze causali da esperimenti controllati. Secondo, le trattazioni filosofiche della causalità probabilistica la includono [Suppes 1970; Reichenbach 1956], e, terzo, i modelli ad equazioni strutturali, che sono forse la classe di modelli causali statistici più ampiamente usata nelle scienze sociali, sono causalmente markoviani. Elaborare le prime due linee esula dagli scopi di questo articolo. Qui cercherò di rendere un po' più esplicito il legame tra i modelli ad equazioni strutturali e l'assunzione markoviana causale.

In un modello ad equazioni strutturali, ciascuna variabile è uguale a una funzione lineare delle sue cause dirette, più un termine d'errore. Dunque il grafo causale della figura 4 si potrebbe tradurre nel seguente modello ad equazioni strutturali:

$$Y = \beta_1 S + \varepsilon_y$$

$$C = \beta_2 S + \varepsilon_c$$

dove  $\beta_1$  e  $\beta_2$  sono coefficienti con valori reali e  $\varepsilon_y$  e  $\varepsilon_c$  sono termini d'errore con varianza strettamente positiva. Se il sistema nella figura 4 è com-

pleto come specificato, ovvero se il grafo causale è completo rispetto alle cause comuni, allora chi elabora modelli ad equazioni strutturali assumerà che  $\varepsilon_y$  e  $\varepsilon_c$  sono indipendenti l'uno dall'altro e da  $S$ . Nei modelli ad equazioni strutturali in cui sono incluse tutte le cause comuni, si assume che i termini d'errore siano indipendenti. Risulta che modelli di tal genere *soddisfano necessariamente* l'assunzione markoviana causale.

La connessione tra i modelli ad equazioni strutturali e la causalità (in quanto questa comporta la risposta di un sistema di interventi) sorge attraverso la connessione tra termini d'errore indipendenti e manipolazioni ideali. Sebbene le manipolazioni ideali forniscano la base semantica per le asserzioni causali, tali manipolazioni sono talvolta solo ideali, e non possono essere realizzate in pratica. Ad esempio, sebbene la povertà possa causare criminalità, non possiamo, per ragioni etiche, compiere degli interventi volti ad impoverire le persone. In situazioni di questo tipo ci dobbiamo limitare a raccogliere dati passivamente. Dal momento che la scienza sperimentale è specializzata nel creare combinazioni nelle quali manipolazioni ideali esistono e sono soggette alle nostra volontà, non bisogna stupirsi se i critici dell'inferenza causale da dati statistici insistono nel dire che gli esperimenti sono l'unico mezzo per formulare asserzioni causali. Ma, a parte il "non posso immaginare in che altro modo si potrebbe fare", qual è la loro argomentazione?

In primo luogo, ci possono essere fonti idealmente selettive di variazioni che esistono in natura, ma che noi non possiamo controllare, né possiamo sperare di farlo un giorno. Ad esempio, la posizione della luna esercita un effetto diretto sul campo gravitazionale degli oceani che causa le maree. Benché la luna sia una fonte che non possiamo controllare, almeno possiamo misurarne alcune caratteristiche.

In altri sistemi, tali fonti ideali di variazioni possono esistere, ma essere fuori dal nostro controllo e inosservabili. In verità, un'interpretazione naturale dei termini d'errore nei modelli ad equazioni strutturali assegna loro precisamente queste proprietà. C'è un unico termine d'errore  $\varepsilon_x$  per ciascuna variabile specificata  $X$ , e in sistemi che includono tutte le cause comuni si assume che  $\varepsilon_x$  sia una fonte di variazione di  $X$  che influenza *direttamente* solo  $X$ . E, benché non possiamo misurarli o controllarli, coloro che elaborano modelli ad equazioni strutturali assumono che tali termini d'errore esistono. È questa assunzione che rende questi sistemi causalmente markoviani.

C'è la tentazione di pensare che anche se interpretiamo i termini d'errore come rappresentazioni di manipolazioni inosservabili e ideali, siamo del tutto bloccati ai fini dell'inferenza causale, semplicemente perché non possiamo osservarli. Ma questa è una fallacia. È vero che non possiamo imparare in merito alla struttura causale di questi sistemi tanto quanto potremmo fare se i termini d'errore fossero osservabili (il mondo degli sperimentatori), ma non segue affatto che non possiamo imparare nulla a tal proposito. In realtà è proprio qui che comincia l'inferenza causale, con sistemi che si assume siano causalmente markoviani.



#### 4. L'inferenza

Accettata l'assunzione markoviana causale, mi rivolgo ora al tema dell'inferenza: passare da dati statistici a conclusioni in merito alla struttura causale. Partendo da dati statistici e conoscenza di sfondo, vogliamo trovare tutte le strutture causali che possono aver generato questi dati. Minore è il numero di assunzioni che facciamo nel limitare la classe di strutture causali possibili, più debole sarà ciò che riusciremo a ottenere inferenzialmente. Bisogna notare, in ogni caso, che non è compito della nostra teoria dettare quali assunzioni un ricercatore dovrà fare, ma solo caratterizzare che cosa si può imparare del mondo, e che cosa non si può imparare, date le particolari assunzioni scelte.

In questo paragrafo discuto alcune delle assunzioni che abbiamo studiato. Ce ne sono molte altre che abbiamo intenzione di studiare in futuro, o che sarebbe interessante studiare. Una classe molto ampia di problemi che non tratterò affatto concerne le inferenze statistiche in merito all'indipendenza: inferire l'insieme di relazioni di indipendenza in una popolazione a partire da un campione. In quanto segue assumerò che i dati sono statisticamente ideali e che di fatto la popolazione è presente davanti a noi, cosicché qualunque asserzione di indipendenza probabilistica possa essere stabilita in modo perfettamente affidabile.

##### 4.1 Fedeltà

La prima assunzione che discuterò è la quella di *fedeltà*. Assumendo che un grafo causale è causalmente markoviano, assumiamo che qualunque popolazione prodotta da questo grafo causale presenta le relazioni di indipendenza ottenute applicando la d-separazione. Non segue, tuttavia, che la popolazione abbia esattamente queste indipendenze, e non altre. Ad esempio, supponiamo che la figura 8 sia un grafo causale che descrive correttamente le relazioni tra esercizio fisico, abitudine al fumo e salute, in cui i segni + e - indicano, rispettivamente, le relazioni positive e inibenti.

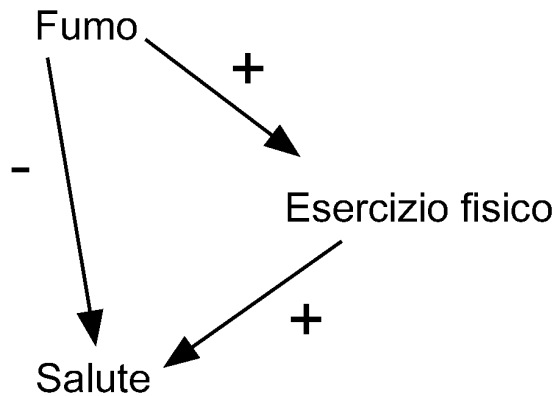


Fig. 8

In questo caso, l'assunzione markoviana causale da sola non pone restrizioni sulle distribuzioni che la struttura potrebbe produrre, perché non otteniamo alcuna indipendenza applicando la d-separazione o la condizione markoviana al suo DAG. Ma in alcune delle distribuzioni che questa struttura potrebbe produrre il fumo potrebbe essere indipendente dalla salute "per coincidenza". Se il fumo ha un effetto negativo diretto sulla salute, ma il fumo ha un effetto positivo sull'esercizio fisico (per quanto assurdo possa sembrare) e l'esercizio ha un effetto positivo sulla salute, allora il fumo serve per inibire direttamente la salute, e per aumentarla indirettamente. Se accade che i due effetti si bilancino perfettamente e quindi si annullino reciprocamente, allora ci potrebbe non essere alcuna associazione tra il fumo e la salute. In un caso di questo tipo possiamo dire che la popolazione è *infedele* al grafo causale che l'ha generata.

Se ci sono relazioni di indipendenza nella popolazione che non sono una conseguenza della condizione markoviana causale (o della d-separazione), allora la popolazione è *infedele*. Assumendo la fedeltà evitiamo di considerare tutti questi casi. Anche se in prima battuta sembra un'assunzione pesante, non lo è. Assumere che una popolazione è fedele significa assumere che qualunque indipendenza si presenta in essa deriva non da un'incredibile coincidenza, ma piuttosto dalla struttura. Una qualche forma di questa assunzione è adottata in ogni scienza. Quando una teoria non può spiegare una regolarità empirica se non invocando una speciale parametrizzazione, gli scienziati non sono soddisfatti dalla teoria, e cercano un'alternativa che possa spiegare la stessa regolarità facendo appello alla struttura anziché alla fortuna. Nel caso dei modelli causali, le regolarità sono relazioni di indipendenza (condizionali), e l'assunzione di fedeltà è solo una codificazione molto chiara della preferenza per modelli che spieghino queste regolarità invocando la struttura e non la fortuna. Questa non è in alcun modo una garanzia; la natura può essere capricciosa. Ma l'esistenza di casi in cui una procedura che assume la fedeltà fallisce sembra un argomento terribilmente debole contro la possibilità di inferenza causale. Ciononostante, i critici continuano a creare casi infedeli e ad esibirli. Assumere la fedeltà sembra ragionevole, ed è un'opzione ampiamente abbracciata dagli scienziati. Il vantaggio inferenziale che si guadagna da questa assunzione nell'inferenza causale è enorme. Senza, tutto quello che possiamo dire sulla base dei dati di indipendenza è che, qualsiasi struttura causale abbia generato i dati, non può implicare alcuna relazione di indipendenza per d-separazione che non sia presente nella popolazione. Con l'assunzione, possiamo dire che, qualsiasi struttura abbia generato i dati, implica per d-separazione *esattamente* le relazioni di indipendenza che sono presenti nella popolazione. Ad esempio, supponiamo di avere una popolazione che coinvolge tre variabili,  $X_1$ ,  $X_2$ ,  $X_3$ , e supponiamo che le relazioni di indipendenza in questa popolazione siano quelle rappresentate nella tavola 1.

TUTTE LE POSSIBILI INDIPENDENZE TRA $X_1, X_2, X_3$	NELLA POPOLAZIONE	NON NELLA POPOLAZIONE
$X_1 \perp\!\!\!\perp X_2$	√	
$X_1 \perp\!\!\!\perp X_3$		√
$X_2 \perp\!\!\!\perp X_3$		√
$X_1 \perp\!\!\!\perp X_2 \mid X_3$		√
$X_1 \perp\!\!\!\perp X_3 \mid X_2$		√
$X_2 \perp\!\!\!\perp X_3 \mid X_1$		√

Tav. 1

Anche se assumiamo che tutti i grafi causalmente markoviani che possono aver prodotto dati con queste relazioni di indipendenza coinvolgono solo  $X_1, X_2$  e  $X_3$ , ci sono ancora nove grafi di tale genere. La sola caratteristica che condividono è che ciascuno ha una qualche relazione diretta tra  $X_1$  e  $X_2$  e tra  $X_2$  e  $X_3$ . Aggiungere la fedeltà riduce l'insieme di nove grafi a un singoletto (vedi figura 9).

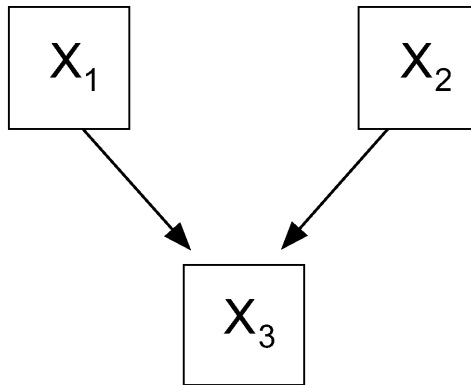


Fig. 9

#### 4.2 Sufficienza causale

In questo esempio siamo riusciti a inferire che sia  $X_1$  che  $X_2$  sono cause dirette di  $X_3$  da una singola indipendenza marginale tra  $X_1$  e  $X_2$ . Questo dà a molti una tregua, com'è giusto. In questo caso abbiamo raggiunto una tale, enorme, leva inferenziale, non solo assumendo la fedeltà, ma anche assumendo la sufficienza causale, che ho rivelato sopra scrivendo: "tutti i grafi causalmente markoviani ... coinvolgono solo  $X_1, X_2$  e  $X_3$ ".

L'assunzione della sufficienza causale è soddisfatta se abbiamo *misurato* tutte le cause comuni delle variabili misurate. Anche se sembra piuttosto simile alle assunzioni in merito alla completezza dei grafi causali, non è

esattamente la stessa cosa. Quando assumiamo che un grafo causale è completo rispetto alle cause comuni, è allo scopo di essere chiari in merito a quali tipi di sistemi stiamo rappresentando con quei grafi. Nel caso dell'inferenza, stiamo facendo due assunzioni: una comporta l'esistenza di qualche grafo causale che sia completo rispetto alle cause comuni e che sia causalmente markoviano; l'altra è un'assunzione in merito alle variabili che abbiamo misurato in contrapposizione a quelle che non abbiamo misurato. Ad esempio, possiamo costruire un modello nel quale specifichiamo una variabile chiamata "Intelligenza" che non possiamo misurare in modo diretto o perfetto, mentre possiamo misurare quattro dei suoi effetti, ad esempio i punteggi di test  $X_1 - X_4$  (figura 10). Supponendo che il grafo causale che rappresenta la variabile Intelligenza e  $X_1 - X_4$  sia completo rispetto alle cause comuni, e che sia causalmente markoviano e fedele a qualsiasi popolazione produce su {Intelligenza,  $X_1 - X_4$ }, allora in questa popolazione varrà la seguente lista di relazioni di indipendenza:

$X_1 \perp\!\!\!\perp X_2 \mid \text{Intelligenza}$   
 $X_1 \perp\!\!\!\perp X_3 \mid \text{Intelligenza}$   
 $X_1 \perp\!\!\!\perp X_4 \mid \text{Intelligenza}$   
 $X_2 \perp\!\!\!\perp X_3 \mid \text{Intelligenza}$   
 $X_2 \perp\!\!\!\perp X_4 \mid \text{Intelligenza}$   
 $X_3 \perp\!\!\!\perp X_4 \mid \text{Intelligenza}$

Dal momento che l'Intelligenza non è misurata, tuttavia, i nostri dati includeranno solo indipendenze che non la coinvolgono (in questo caso, è l'insieme vuoto). Dunque il grafo causale che concerne la variabile Intelligenza e  $X_1 - X_4$  è completo rispetto alle cause comuni, ma le variabili misurate,  $X_1 - X_4$ , non sono causalmente sufficienti. Per riassumere, l'assunzione markoviana causale, benché includa una forma rappresentativa di sufficienza causale, è un'assunzione in merito al modo in cui causalità e probabilità sono connesse, mentre la sufficienza causale è un'assunzione in merito a quanto siamo riusciti a misurare. Ho finora discusso tre diverse assunzioni:

- 1) l'assunzione markoviana causale: se si specifica accuratamente un grafo causale  $G$  su un insieme di variabili  $V$  (in cui  $V$  include tutte le cause comuni di coppie in  $V$ ), *almeno* le relazioni di indipendenza ottenute applicando la d-separazione a  $G$  valgono nella distribuzione di probabilità su  $V$ ;
- 2) l'assunzione di fedeltà: *esattamente* le relazioni di indipendenza ottenute applicando la d-separazione a  $G$  valgono nella distribuzione di probabilità su  $V$ ;
- 3) l'assunzione della sufficienza causale: l'insieme di variabili *misurate*  $M$  include tutte le cause comuni di coppie in  $M$ .

Nell'esempio concernente la fedeltà, siamo riusciti a inferire la struttura causale unica nella figura 9 dalla singola indipendenza marginale  $X_1 \perp\!\!\!\perp X_2$  facendo tutte e tre le assunzioni. È possibile fare inferenze in merito alla/e struttura/e sottese ai dati senza l'assunzione della sufficienza causale, ma, ovviamente, non si apprende altrettanto.

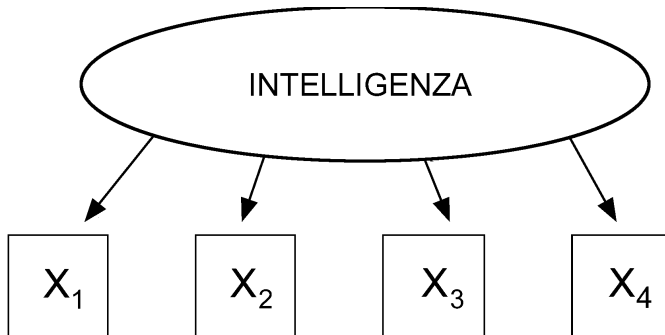


Fig. 10

Quando non assumiamo la sufficienza causale, assumiamo ugualmente che *ci sia* qualche struttura che include le variabili misurate (e forse altre variabili), che è completa rispetto alle cause comuni e che soddisfa l'assunzione markoviana causale, ma dobbiamo riconoscere che potremmo non aver misurato tutte le cause comuni. Pertanto, qualunque sia l'algoritmo che usiamo per passare da relazioni di indipendenza a tutti i grafi causali che possono aver prodotto tali relazioni di indipendenza, l'insieme dei grafi deve includere membri che hanno cause comuni che non abbiamo misurato. Nell'esempio nella figura 9 abbiamo misurato  $X_1 - X_3$ , e abbiamo osservato una sola indipendenza:  $X_1 \perp\!\!\!\perp X_2$ . Se assumiamo la condizione markoviana causale e la fedeltà, ma non la sufficienza causale, l'insieme dei dieci grafi causali che produrrebbe esattamente questa indipendenza è quello raffigurato nella figura 11, in cui le variabili  $T$  nei cerchi sono le cause comuni che potremmo non aver misurato.

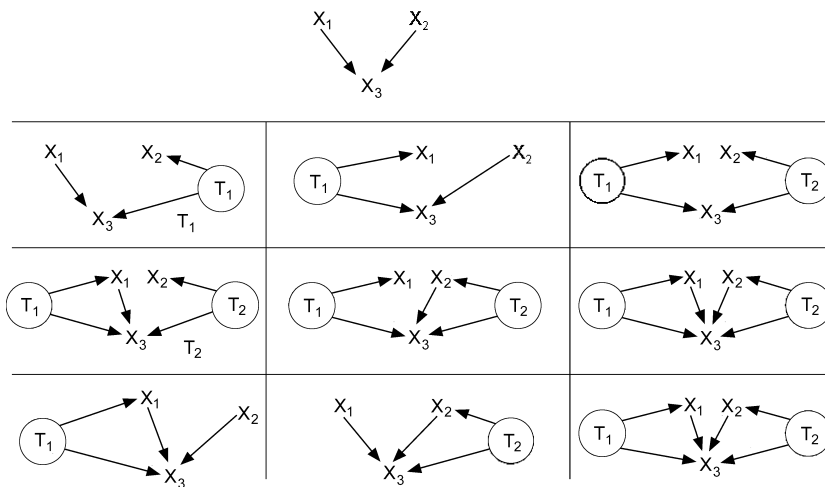


Fig. 11

A dire il vero, questo insieme è ancora troppo piccolo, poiché lì dove noi abbiamo specificato una sola causa comune di due variabili (quali  $X_1$  e  $X_3$ ) non misurata, chiamata " $T_1$ ", ci potrebbe in realtà essere qualunque numero di cause comuni distinte, non misurate, di  $X_1$  e  $X_3$ . Dunque, ogniqualvolta  $T_1$  compare come una causa comune (ad esempio, di  $X_1$  e  $X_3$ ), è in realtà un'abbreviazione per: esistono alcune cause comuni non misurate di  $X_1$  e  $X_3$ .

Anche se lasciar cadere l'assunzione della sufficienza causale ha ridotto considerevolmente il nostro potere inferenziale, non lo ha completamente eliminato. Si noti che in nessuna delle strutture nella figura 11  $X_3$  è una causa di qualche altra variabile. Così abbiamo appreso qualcosa in merito a quali relazioni causali non sussistono:  $X_3$  non è una causa di  $X_1$  o di  $X_2$ , pur essendo associata ad entrambe. In altre parole, abbiamo inferito dai dati di indipendenza quanto segue: se dovessimo idealmente manipolare  $X_3$ , non faremmo nulla per alterare  $X_1$  o  $X_2$ .

Possiamo mai acquisire conoscenza concernente quali relazioni causali esistono senza assumere la sufficienza causale? Sì, purché misuriamo almeno quattro variabili oppure facciamo assunzioni aggiuntive. Ad esempio, se vengono osservate le seguenti indipendenze tra  $X_1 - X_4$ , allora, assumendo la condizione markoviana causale e la fedeltà, possiamo concludere che in ogni grafo che può avere generato questi dati  $X_3$  è una causa di  $X_4$ .

$$X_1 \perp\!\!\!\perp X_2$$

$$X_1 \perp\!\!\!\perp X_4 \mid X_3$$

$$X_2 \perp\!\!\!\perp X_4 \mid X_3$$

Ovvero: se la condizione markoviana causale e la fedeltà sono soddisfatte, allora da queste relazioni di indipendenza possiamo concludere che una manipolazione di  $X_3$  farebbe variare la probabilità di  $X_4$ . Aggiungere altri tipi di conoscenza, ad esempio conoscenza in merito all'ordine temporale in cui si presentano le variabili, spesso migliora la situazione.

### 5. Conclusioni

Contrariamente a quello che alcuni ritengono essere il nostro scopo, non intendiamo fare uscire miracolosamente conigli causali da un cappello statistico. La nostra teoria dell'inferenza causale si interroga su che cosa può e che cosa non può essere appreso in merito alla struttura causale a partire da un insieme di assunzioni che pare vengano comunemente fatte nella pratica scientifica. È quindi una teoria concernente l'effetto inferenziale di una varietà di assunzioni molto più di quanto non sia la difesa di particolari assunzioni. Ci sono situazioni nelle quali è irragionevole sostenere l'assunzione markoviana causale (ad esempio, nella meccanica quantistica), la sufficienza causale sembra raramente ragionevole, e ci sono certe situazioni nelle quali si può non voler assumere la fedeltà (ad esempio, se alcune variabili sono completamente determinate da altre).

Se il nostro lavoro ha un qualche impatto positivo in pratica, sarà tanto nel gettare dubbi sulle teorie di maggior successo rendendo semplice mostrare che esistono alternative equivalenti ragionevoli, quanto nell'estrarre conclusioni causali da dati statistici. Se la nostra posizione riesce a chiarire il razionale scientifico che sottende all'inferenza causale, cosa che è il nostro obiettivo reale, allora il suo risultato più importante sarà quello di cambiare il modo in cui le ricerche vengono concepite e i dati raccolti.