UNIVERSITA' DEGLI STUDI DI PADOVA

Facoltà di Scienze Statistiche

Corso di Laurea Specialistica

in

Scienze Statistiche Economiche Finanziarie e Aziendali

Tesi di Laurea

Modelli strutturali con variabili categoriali: un approccio a classi latenti

Relatore: Prof. Adriano Paggiaro

Laureanda: Gloria Gheno

Matricola 545547

ANNO ACCADEMICO 2010/2011

Sommario

Introduzione

Molti dei dati utilizzati per studiare fenomeni economici, politici e sociali sono raccolti con survey, in cui i soggetti intervistati esprimono preferenze, opinioni e informazioni, e per tale ragione hanno natura categoriale. I dati siffatti possono essere analizzati con tecniche statistiche proprie, o con metodologie per dataset continui, ipotizzando una loro approssimazione.

Una tecnica ormai consolidata per i dati categoriali è quella che si fonda sui modelli log-lineari causali, la cui divulgazione si deve ai paper di Goodman degli inizi degli anni '70. In particolare, Goodman (1973) introduce la causalità delle relazioni tra le variabili riuscendo a formulare un modello simile concettualmente a quelli già in uso per le variabili continue, gli structural equation models (SEM).

I SEM, la cui definizione si può far risalire già agli anni venti (Wright 1921), permettono sia di verificare teorie, con la modellazione confermativa, sia di svilupparne, con quella esplorativa. Uno dei punti di forza di tale metodologia è rappresentato dall'analisi delle variabili latenti, le quali non sono osservate direttamente nel dataset, ma ottenibili da questo mediante opportune trasformazioni. Questa ricerca di semplificazione di un dataset, ipotizzando che il legame (misurato statisticamente dalla covarianza) tra le sue variabili sia dovuto a variabili non osservate, trova la sua più nota espressione nell'analisi fattoriale.

L'importanza delle variabili latenti nelle ricerche sociali può essere ricondotta al problema che molti fattori, analizzati da questa tipologia di studi, non sono osservabili direttamente, ma solo attraverso loro manifestazioni. Goodman, riconoscendo l'importanza delle variabili latenti, le introduce nel 1974 nella metodologia log-lineare con un'analisi fattoriale che è concettualmente un ibrido tra quella confermativa e quella esplorativa SEM. Per avere un'analisi fattoriale esplorativa propriamente detta bisogna aspettare il 2001, con il paper di Vermunt e Magidson.

Le molteplici applicazioni della metodologia SEM, comunque, non sono solo dovute alla sua possibilità di studiare i legami tra variabili osservate e variabili latenti, ma anche alla sua capacità di spiegare causalmente le relazioni tra quest'ultime. Il modello SEM, nella sua più generale formulazione, è composto da una parte di misura, in cui si raggruppano le variabili osservate in variabili latenti, e da una parte strutturale, in cui si esaminano le relazioni causali tra le variabili latenti determinate nella parte precedente. L'analisi fattoriale, che prevede solo la parte di misura, è quindi un caso specifico del SEM, come lo è il SEM a variabili osservate che considera solo la parte strutturale. Bergsma, Croon e Hagenaars (2009) creano il modello log-lineare causale a classi latenti: un modello concettualmente identico al SEM per i dati categoriali. Come il SEM, infatti, anche il modello log-lineare causale a classi latenti è suddiviso in una parte di misura ed in una parte strutturale. All'interno di questa ricerca di formulare un modello SEM per i dati categoriali si inserisce questa tesi.

Nel primo capitolo si descrivono il modello log-lineare e la sua evoluzione causale con variabili osservate. L'identificazione e l'interpretazione dei relativi parametri sono poi analizzate con dati creati ad hoc senza alcuna variabilità casuale. I dati dimostrano

quanto siano differenti le 2 modellazioni sebbene le variabili siano costruite partendo dagli stessi valori dei parametri, che risultano pertanto interpretativamente diversi: quelli log-lineari “tradizionali” sono una misura analoga alla covarianza nei modelli SEM, mentre quelli “causali” sono relativi all’effetto diretto. Stimare un modello creato causalmente con la metodologia “tradizionale” (chiamata anche ABC per ricordare che il legame tra le 3 variabili è dovuto esclusivamente alla covarianza) porta quindi all’identificazione erronea dei parametri e viceversa. Una successiva analisi più dettagliata mostra che nel modello log-lineare causale si possono ricreare partendo da un modello completo molte tipologie di sub-modelli presenti nella letteratura causalistica. Partendo da un modello ricorsivo completo con 3 variabili si ottengono, eliminando parametri, un modello con solo l’effetto indiretto e due differenti casi con solo l’effetto diretto.
Per uguagliare esattamente un modello SEM, il modello log-lineare richiederebbe che gli effetti causali fossero tutti esplicitati, fatto che non avviene, in quanto in letteratura si è determinato, con Goodman (1973) e successivamente con Hagenaars (1990), solo l’effetto diretto. Un passo avanti su questo punto si può trovare nel capitolo 6 di Croon et al (2009), in cui si asserisce che l’effetto totale di una variabile A su una variabile C può essere analizzato nella tabella di frequenza AC e quindi la sua nullità richiede che A e C siano indipendenti se si studiano dalla medesima tabella.
Un contributo innovativo di questa tesi è dato da una formula per calcolare l’effetto totale in un modello a 3 variabili dicotomiche direttamente dai parametri che risultano dalla stima della tabella di frequenza ABC. Rifacendosi alla letteratura causale, tra cui Pearl (2000 e 2001) e Rao,Li e Roth (2008), si scompone l’effetto totale in effetto diretto e indiretto come avviene nella classica delle metodologie. L’effetto indiretto così calcolato, cioè come differenza tra effetto totale ed effetto diretto, deve però essere successivamente suddiviso in effetto indiretto propriamente detto ed effetto cella, che deriva dalla presenza di più effetti diretti influenti su una medesima variabile. Si definisce effetto indiretto totale quello calcolato residualmente e indiretto quello tipico presente in letteratura. Le dimostrazioni di come sono stati ricavati questi effetti sono presentate in Appendice, mentre la loro applicazione si osserva nei sub-modelli, nuovamente generati da dati in cui non è presente la variabilità così da permettere l’esatta identificazione dei parametri.
Nel capitolo 2 si descrivono le caratteristiche principali del modello SEM, mostrando inoltre che l’effetto totale in un modello a 3 variabili osservate A,B, C si può calcolare, similmente alla metodologia log-lineare causale, dalla matrice di varianza-covarianza AC anziché dai parametri ricavati dalla matrice di varianza-covarianza ABC.
Nel capitolo 3 si confronta la metodologia log-lineare causale con quella SEM, dopo aver introdotto le variabili latenti nei modelli log-lineari così da poter confrontare in maniera esaustiva il modello SEM tradizionale con un modello concettualmente simile per i dati categoriali, il modello log-lineare causale con variabili latenti (Croon e al. 2009). Questo approccio prevede come per il SEM una parte di misura ed una parte strutturale, classica ripartizione di un modello structural equation model.
Per possedere tutti gli strumenti per un effettivo confronto si necessita di un’analisi fattoriale a classi latenti, introdotto rifacendosi a Magidson e Vermunt (2001) e

mostrandone la procedura attraverso 2 dataset, il cui processo di generazione è ancora una volta privo di casualità. Il primo, composto da 5 variabili (A,B,C,D,E), serve a dimostrare come da un dataset "completo",quindi con tutte le variabili osservate, si passa ad un dataset in cui il legame tra le variabili è dovuto alla presenza di una variabile latente. Si ipotizza che A influisca direttamente su B, C, D ed E e che il legame tra queste ultime 4 variabili sia dovuto esclusivamente ad A, ipotesi chiamata locale indipendenza e che sta alla base del latent class model. Per spiegare la procedura di un'analisi fattoriale si considera A latente e B, C, D ed E osservate, per tale ragione il dataset che si va a stimare è composto dalla tabella di frequenza di BCDE. Il secondo dataset è creato con la stessa procedura del primo ma differenziandosi per il numero di variabili che aumenta a 6 e per la relazione che c'è tra di esse. Continuando ad utilizzare la locale indipendenza, si suppone che A influisca su B e C ed F influisca su D ed E; A e F sono tra loro correlate e si ipotizzano latenti così il dataset su cui si svolge la factor analysis è composto ancora dalla tabella di frequenza BCDE. Il modello a 2 fattori, così denominati dalla procedura SEM invece che variabili latenti, è quello scelto dall'analisi, coerentemente con la costruzione dei dati.

La parte strutturale è quella che determina propriamente le relazioni nella loro formulazione causale e quindi è la più importante quando si cerca di capire la causalità degli eventi. Ogni ricercatore, infatti, ricerca i fattori che influiscono sul fenomeno oggetto di studio e solo con l'analisi strutturale riesce a determinare sia l'intensità sia la direzione del legame che unisce i vari elementi sottoposti ad esame. Per questo motivo ci si sofferma sugli effetti causali, paragonando il modello SEM a variabili osservate a quello log-lineare causale utilizzando i dataset del capitolo 1, rappresentanti i sub-modelli ottenibili da un modello completo. I dati sono quindi costruiti a partire dai parametri log-lineari e dalla loro relativa tabella di frequenza e successivamente trasformati in matrice di varianza-covarianza affinché sia possibile un'analisi SEM.

Dal confronto delle stime dei dataset si rileva che le 2 metodologie portano sostanzialmente ai medesimi risultati. L'assenza delle latenti in quest'analisi non è importante poiché quanto assunto per i modelli a variabili osservate può essere esteso a modelli con variabili latenti, naturalmente con le dovute attenzioni: l'analisi fattoriale, infatti, può portare a variabili latenti con scale differenti nelle 2 metodologie e ciò deve essere considerato quando si paragonano gli effetti strutturali. La scelta di un'analisi con i soli dati osservati è stata fatta per rendere più agevole e diretto il confronto. Da questo studio si conclude che i modelli SEM e quelli log-lineari causali a classi latenti forniscono simili procedure di analisi ed interpretazione dei risultati.

Il modello log-lineare causale, comunque, "soffre" la crescita della dimensione, dovuta all'aumentare delle variabili e/o delle loro classi, della tabella di frequenza multivariata: un dataset troppo "grande" può portare a problemi di identificazione dei parametri ed a modelli instabili.

Nell'ultimo capitolo si confrontano le 2 metodologie in un caso pratico, utilizzando i dati di un questionario distribuito durante la festa del prosciutto di SAURIS (UD) che si svolge nel secondo e terzo fine settimana di Luglio. Questa survey si colloca quindi nell'analisi del turismo enogastronomico e nella commercializzazione dei prodotti tipici: sempre più spesso, infatti, piccoli borghi ritrovano l'antica vitalità e quindi la

loro sopravvivenza nel turismo e nell'artigianato. Con l'analisi dei dati raccolti si cercano di capire i fattori che determinano l'affluenza alla manifestazione e la vendita del prodotto.
Si colloca in questo "scenario" il confronto fra la metodologia SEM e quella log-lineare causale a classi latenti, paragone che avviene in alcuni casi studio di modelli ristretti a poche variabili osservate, senza alcuna pretesa di un esaustivo analisi di marketing sul turismo enogastronomico. Centrale nello studio è il concetto di "soddisfazione del cliente", tipico nell'analisi di marketing e suo obiettivo strategico: aumentando la soddisfazione, cioè eccedendo le aspettative del cliente (Anderson e Mitall 2000), si guadagnano fette di mercato e quindi maggiori introiti. Il cliente-turista più soddisfatto, infatti, ritornerà nella località turistica e la consiglierà ad altri, comportamenti che in marketing vengono definiti behavioral intentions. La relazione tra soddisfazione del cliente e behavioral intentions è analizzata in molteplici studi (per esempio Chen & Tsai 2009, Cronin et al 2000, Baker& Crompton 2000).
Nel caso specifico di Sauris si studia come la soddisfazione sulla festa influisca sul comportamento futuro nei confronti della festa e del prosciutto omonimi con la metodologia SEM e con quella log-lineare causale a classi latenti. Nel primo caso, utilizzando 8 variabili, che esprimono altrettanti impressioni sulla festa di Sauris, si confronta un'analisi fattoriale esplorativa nella metodologia SEM e in quella a classi latenti. Il risultato è uguale in entrambi i modelli e dimostra che le 8 variabili sono indicatori di una sola variabile latente.
Nel secondo caso si analizza un dataset formato da 6 variabili osservate indicatrici delle 3 variabili latenti seguenti: "interesse verso il prosciutto" (X), "soddisfazione della festa"(K), "comportamento futuro nei confronti del prosciutto"(V) . Entrambe le metodologie portano alla scelta di un modello completo in cui X influisce direttamente sia su K che su V e K influisce direttamente su V e che può essere così schematizzata↑X=>↑K=>↑ V e ↑X=>↑V. L'effetto totale e l'indiretto , calcolati nella metodologia log-lineare con la formulazione proposta nel capitolo 1, forniscono per entrambi i metodi che un ↑X=>↑V.
Nel terzo caso si raffronta un dataset formato da 6 variabili osservate indicatrici delle 3 variabili latenti seguenti: "interesse verso il prosciutto" (X), "soddisfazione della festa"(K), "comportamento futuro nei confronti della festa"(H) . Il modello che ne deriva non si adatta bene ai dati, ma comunque si sceglie di inserirlo in questa tesi in quanto dimostra che entrambe le metodologie preferiscono lo stesso sub-modello a quello completo. Un'altra particolarità di questo processo di stima è che l'analisi fattoriale, portando a variabili con scale differenti nelle 2 metodologie, influisce sul "segno apparente" degli effetti. Nel modello SEM X, K e V sono variabili crescenti, in LCA X e H sono crescenti e K è decrescente e ciò porta ad una diversa interpretazione dei parametri strutturali: un effetto diretto positivo di X su K in SEM equivale ad un effetto negativo in LCA, ma il sub-modello scelto da entrambe le metodologie mostra la seguente relazione sostanziale in termini di effetti diretti: ↑X=>↑K=>↑ H.
Il capitolo si conclude con un modello completo che è particolarmente interessante per vedere il ruolo dell'effetto cella, "interazione" insita solo nei modelli log-lineari. Si utilizza un dataset formato da 6 variabili osservate indicatrici delle 3 variabili latenti

seguenti: “qualità della festa” (W), “soddisfazione della festa”(K), “comportamento futuro nei confronti del prosciutto”(K). Entrambe le metodologie portano alla scelta del medesimo modello, che in termini interpretativi degli effetti diretti può essere riassunto in $\uparrow$W=>$\uparrow$K=>$\uparrow$ V e $\uparrow$W=>$\uparrow$ V. L’effetto totale di W su V produce una relazione interpretazionale positiva nelle 2 metodologie ma l’effetto cella attenua l’effetto indiretto portando il valore numerico dell’effetto totale ad essere molto simile a quello dell’effetto diretto.

Questo conferma l’intuizione nata all’inizio del confronto: i modelli SEM e quelli log-lineari causali forniscono un medesimo modello causale in termini di interpretazione sostanziale, ma le intensità delle relazioni causali non sono confrontabili. Questa affermazione ha una sua logica in quanto i parametri log-lineari misurano l’influenza di un particolare fattore sulla probabilità di essere nella classe i della variabile A, quelli SEM misurano l’influenza del fattore proprio sulla classe i della variabile A.

Capitolo 1:
modelli causali con variabili categoriali

1.1 I MODELLI LOG-LINEARI

1.1.1 Variabili categoriali

I modelli log- lineari sono usati per l'analisi dei dati categoriali, cioè di dataset composti da variabili osservate di tipo categoriale, caratterizzate da un numero finito di categorie. Le variabili categoriali assumono il nome di dicotomiche se le categorie sono 2, di politomiche se sono in numero maggiore di 2.

Le variabili categoriali si suddividono in nominali, ordinali e intervallari. Le variabili sono nominali se composte da due o più categorie non ordinate. Un esempio di variabile nominale dicotomica, cioè a due categorie, è la variabile sesso: le 2 modalità, maschio o femmina, non sono una più importante dell'altra. Una variabile nominale a tre categorie è il colore degli occhi: azzurri, verdi e neri. In questo caso non esiste un ordine: avere gli occhi azzurri non è più o meno importante di avere gli occhi neri.

Le variabili ordinali si differenziano dalle variabili nominali perché le categorie sono ordinate. Un esempio di variabile ordinale a 4 categorie è il livello di istruzione: diploma elementare, diploma di scuola media inferiore, diploma di scuola media superiore e diploma di laurea. In questo caso le categorie hanno un ordine: per esempio, avere un diploma di scuola media superiore vuol dire avere un livello di istruzione maggiore rispetto a chi ha un diploma di scuola media inferiore.

Le variabili intervallari sono simili alle ordinali, la sola differenza è che le categorie ordinate sono rappresentate da classi in cui è stata suddivisa una variabile continua. Un esempio di variabile intervallare è la suddivisione della popolazione di una regione basata su classi di età. La prima classe rappresenta il numero di abitanti con un'età compresa tra 0-20, la seconda tra 21-40, la terza tra 41-60, la quarta tra 61-80, la quinta oltre gli 80. In questo esempio la variabile sottostante, cioè l'età, è continua e le 5 classi in cui è raggruppata sono ordinate. La prima classe rappresenta la popolazione più giovane e la quinta quella più vecchia.

Ci possono essere anche delle variabili che sono "tra" le ordinate e le intervallari (per esempio una variabile a 5 punti di una scala Likert con categorie uguali a: "concordo pienamente", "concordo", "neutrale", "disaccordo", "disaccordo pienamente") in quanto si immagina che ci sia una variabile sottostante continua, ma che non si conoscano i valori degli intervalli.

1.1.2 Distribuzioni congiunte e marginali

Per studiare le relazioni tra le variabili categoriali si analizza la tabella di frequenza multivariata. Nel modello più semplice, a sole 2 variabili A e B dicotomiche, assumenti valori per esempio 1 o 2, la forma è quella in tabella 1

Tabella 1: tabella di frequenza bivariata

A \ B	B=1	B=2	totale
A=1	$F^{A=1\,B=1}$	$F^{A=1\,B=2}$	$F^{A=1}$
A=2	$F^{A=1\,B=2}$	$F^{A=2\,B=2}$	$F^{A=2}$
totale	$F^{B=1}$	$F^{B=2}$	N

In questo caso non essendo precisato che tipo di variabili categoriali siano A e B , 1 e 2 rappresentano il nome della categoria. Le F rappresentano le frequenze assolute, cioè il numero di volte in cui quel fenomeno è stato osservato nella popolazione. Per esempio $F^{A=1\,B=1}$ è il numero di volte che le modalità A=1 e B=1 sono state osservate congiuntamente nella popolazione; $F^{A=1}$ rappresenta la frequenza marginale, cioè il numero di volte che nella popolazione si è osservata la modalità A=1 senza tenere conto di quale valore assuma la variabile dicotomica B. Il valore N rappresenta la numerosità della popolazione. Passando alle frequenze relative o probabilità (π) la tabella 1 diviene:

Tabella 2: tabella delle frequenze relative o probabilità

A \ B	B=1	B=2	totale
A=1	$\pi^{A=1\,B=1}$	$\pi^{A=1\,B=2}$	$\pi^{A=1}$
A=2	$\pi^{A=2\,B=1}$	$\pi^{A=2\,B=2}$	$\pi^{A=2}$
totale	$\pi^{B=1}$	$\pi^{B=2}$	1

La relazione tra frequenze (F) e probabilità (π) è

$\pi^{A=i\,B=j} = F^{A=i\,B=j}/N$ con i,j=1,2

$\pi^{A=i} = F^{A=i}/N$ con i=1,2

$\pi^{B=j} = F^{B=j}/N$ con j=1,2

Le probabilità che tengono in considerazione sia A che B, cioè $\pi^{A=i\,B=j}$, sono dette congiunte, quelle che tengono in considerazione solo A o solo B, cioè $\pi^{A=i}$ o $\pi^{B=j}$, sono dette marginali. Esse possono essere ricavate dalle congiunte in questo modo:

$\pi^{A=i} = \sum_B \pi^{A=i\,B=j}$ e $\pi^{B=j} = \sum_A \pi^{A=i\,B=j}$

Per studiare se le variabili A e B sono indipendenti o dipendenti si controlla se la probabilità congiunta ($\pi^{A=i\ B=j}$) può essere decomposta come prodotto delle due marginali ($\pi^{A=i}\ e\ \pi^{B=j}$), cioè

$$\pi^{A=i\ B=j} = \pi^{A=i}\pi^{B=j}$$

Un modello più complesso può essere quello a 3 variabili, A, B, C, per semplicità tutte dicotomiche. La tabella di frequenza multivariata può essere così scritta:

Tabella 3: tabella di frequenza multivariata a 3 variabili

A	B	C	
1	1	1	$F^{A=1\ B=1\ C=1}$
1	1	2	$F^{A=1\ B=1\ C=2}$
1	2	1	$F^{A=1\ B=2\ C=1}$
1	2	2	$F^{A=1\ B=2\ C=2}$
2	1	1	$F^{A=2\ B=1\ C=1}$
2	1	2	$F^{A=2\ B=1\ C=2}$
2	2	1	$F^{A=2\ B=2\ C=1}$
2	2	2	$F^{A=2\ B=2\ C=2}$
			N

Le frequenze $F^{A=i\ B=j\ C=k}$, con i,j,k=1,2, rappresentano il numero di volte in cui nella popolazione si sono osservate congiuntamente le modalità i per la variabile A, j per la variabile B e k per la variabile C. Le frequenze marginali non sono inserite nella tabella per problemi di rappresentazione grafica. Esse sono semplicemente calcolabili dalle frequenze $F^{A=i\ B=j\ C=k}$. Per esempio $F^{A=i}$ è la somma dei seguenti 4 addendi: $F^{A=i\ B=1\ C=1}$ $F^{A=i\ B=1\ C=2}$, $F^{A=i\ B=2\ C=1}$ e $F^{A=i\ B=2\ C=2}$.

Come nel caso a 2 variabili, anche in questo si possono calcolare le probabilità ($\pi^{A=i\ B=j\ C=k}$) dividendo le frequenze ($F^{A=i\ B=j\ C=k}$) per N e ricavando le tre marginali ($\pi^{A=i}, \pi^{B=j}, \pi^{C=k}$) dalle congiunte ($\pi^{A=i\ B=j\ C=k}$):

$$\pi^{A=i} = \sum_C \sum_B \pi^{A=i\ B=j\ C=k}$$

$$\pi^{B=j} = \sum_C \sum_A \pi^{A=i\ B=j\ C=k}$$

$$\pi^{C=k} = \sum_A \sum_B \pi^{A=i\ B=j\ C=k}$$

Nel caso a 3 variabili l'indipendenza non è più solo quella tra A e B, come nel caso a 2 variabili, ma per avere la completa indipendenza si considera C e quindi la probabilità congiunta ABC deve potersi scomporre nelle tre marginali di A, di B e di C, cioè:

$$\pi^{A=i\ B=j\ C=k} = \pi^{A=i}\ \pi^{B=j}\ \pi^{C=k}$$

Si è voluto illustrare un esempio a 3 variabili anche se del tutto uguale a quello a 2 variabili perché c'è la possibilità di avere 2 variabili dipendenti e una indipendente dalle altre 2. Per esempio, se A e B sono dipendenti e C è indipendente da A e B, allora la probabilità congiunta delle tre variabili si scompone:

$$\pi^{A=i\,B=j\,C=k} = \pi^{A=i\,B=j}\pi^{C=k}$$

dove

$$\pi^{A=i\,B=j} = \sum_{C} \pi^{A=i\,B=j\,C=k}$$

I modelli si possono rendere sempre più complessi aumentando il numero di categorie e/o il numero di variabili, ma l'analisi rimane sostanzialmente simile a quella sopra illustrata per le tabelle a 3 variabili. Da qui in poi si analizzerà un modello a 3 variabili (A, B, C) politomiche (A con I categorie, B con J categorie e C con K categorie).

1.1.3 Specificazione del modello

I modelli log-lineari si basano sulla tabella di frequenza multivariata, cioè sulla relazione tra le frequenze della tabella di frequenza multivariata ($F^{A=i\,B=j\,C=k}$) e gli effetti, con nessuna connotazione causale, delle singole variabili ($\tau^{A=i}, \tau^{B=j}, \tau^{C=k}$) e delle loro interazioni ($\tau^{A=iB=j}, \tau^{A=i\,C=k}, \tau^{B=j\,C=k}, \tau^{A=i\,B=j\,C=k}$):

$$F^{A=i\,B=j\,C=k} = \tau\, \tau^{A=i}\tau^{B=j}\tau^{C=k}\tau^{A=iB=j}\tau^{A=i\,C=k}\tau^{B=j\,C=k}\tau^{A=i\,B=j\,C=k} = \mathrm{N}\pi^{A=i\,B=j\,C=k} \quad (1)$$

dove la frequenza $F^{A=i\,B=j\,C=k}$ è uguale alla probabilità congiunta $\pi^{A=i\,B=j\,C=k}$ per N, numerosità della popolazione.

Il termine log- lineare si basa sulla trasformazione logaritmica[1] dell'uguaglianza (1):

$$\log(\mathrm{N}\pi^{A=i\,B=j\,C=k}) =$$

$$\lambda + \lambda^{A=i} + \lambda^{B=j} + \lambda^{C=k} + \lambda^{A=iB=j} + \lambda^{A=i\,C=k} + \lambda^{B=j\,C=k} + \lambda^{A=i\,B=j\,C=k} \quad (2)$$

Se una τ è uguale ad 1 o se un λ è uguale a 0, tale effetto non esiste. Un modello viene definito saturo quando compaiono tutti gli effetti. Se alcuni effetti non sono presenti, il modello si definisce non saturo.
Esistono 2 tipi di modelli non saturi: quello gerarchico e quello non gerarchico. Nel modello gerarchico se un parametro τ è uguale ad 1 (o $\lambda = 0$), allora tutti gli effetti dello stesso ordine o superiore sono uguali ad 1 [per esempio se $\tau^{A=iB=j} = 1 => \tau^{A=i\,C=k} = \tau^{B=b\,C=k} = \tau^{A=i\,B=j\,C=k} = 1$]. Si definiscono non gerarchici i modelli che non hanno le restrizioni che seguono una gerarchia, come quella sopra. La distinzione fra gerarchici e non gerarchici è importante per definire il metodo di stima.

---

[1] In tutta la tesi si usa indistintamente sia la simbologia log sia quella ln o in base al riferimento dell'autore o per rendere chiara la base.

Le costanti τ e λ riflettono il livello medio di tutte le celle di frequenza. Esse riflettono puramente la grandezza del campione. La presenza degli effetti ad una variabile comporta che la frequenza (o la probabilità, poiché il passaggio da frequenza a probabilità non influisce sugli effetti) della cella della categoria presa in esame sia più grande o più piccola rispetto ad ognuna delle altre categorie. In generale questa misura non è molto interessante nell'analisi log-lineare. Per spiegare meglio si può prendere come esempio l'effetto $\lambda^{A=i}$: esso misura quanto più grande o più piccola sia la frequenza con A=i rispetto alle frequenze A≠i,cioè indica quanto più ampia sia la probabilità che si appartenga ad A=i piuttosto che a qualsiasi altra categoria di A.
Gli effetti a 2 variabili (ad esempio $\lambda^{A=iB=j}$) sono una misura dell'associazione tra le 2 variabili (in questo caso A e B). Se A e B sono indipendenti $\lambda^{A=iB=j}$ è uguale a 0 , se A e B sono dipendenti $\lambda^{A=iB=j}$ è diversa da 0. La frequenza $F^{A=i\,B=j\,C=k}$ è influenzata dal fatto di avere A=i congiuntamente al fatto di avere B=j per mezzo di $\lambda^{A=iB=j}$. Gli effetti a 3 variabili mostrano come la frequenza $F^{A=i\,B=j\,C=k}$ devii dalle aspettative formulate in base al valore degli effetti di ordine inferiore ($\lambda^{A=i}$, $\lambda^{B=j} e\ \lambda^{C=k}$; $\lambda^{A=i\,B=j}$, $\lambda^{A=i\,C=k} e\ \lambda^{B=j\,C=k}$). Se si passa dalle frequenze ($F^{A=i\,B=j\,C=k}$) alle probabilità ($\pi^{A=i\,B=j\,C=k}$), la(1) e la (2) divengono:

$$\pi^{A=i\,B=j\,C=k} = \eta\,\tau^{A=i}\tau^{B=j}\tau^{C=k}\tau^{A=iB=j}\tau^{A=i\,C=k}\tau^{B=j\,C=k}\tau^{A=i\,B=j\,C=k}$$

$$\log\pi^{A=i\,B=j\,C=k} = u + \lambda^{A=i} + \lambda^{B=j} + \lambda^{C=k} + \lambda^{A=i\,B=j} + \lambda^{A=i\,C=k} + \lambda^{B=j\,C=k} + \lambda^{A=i\,B=j\,C=k}$$

Tutti i parametri τ e λ continuano ad essere interpretati come per le frequenze. Il termine u, e di conseguenza η, divengono costanti di normalizzazione per assicurare che la somma delle probabilità in tutte le possibili combinazioni di i,j e k sia uguale ad 1.

1.1.4 Identificazione e stima

Dopo aver specificato il modello per le variabili A, B e C (la funzione (1) o la sua trasformazione logaritmica (2)) se ne controlla l'identificazione, verificando se i parametri τ o la loro trasformazione logaritmica λ sono univocamente determinati. Come affermato da Hagenaars (1990) la presenza di più soluzioni per i parametri λ e τ non crea problemi in quanto non si interpreta l'effetto (per esempio $\lambda^{A=i\,B=j}$) ma la relazione esistente tra i vari λ associati alle varie categorie (per esempio la relazione tra $\lambda^{A=iB=j}$ e $\lambda^{A=oB=s}$ con i≠o e j≠s). Questa interpretazione porta ad una soluzione generale del problema dell'identificabilità. Esistono 2 tipi di parametrizzazioni tipicamente utilizzate per l'identificabilità: il dummy code e l'effect code, che sono arbitrarie parametrizzazioni dello stesso modello con le stesse frequenze attese stimate. I parametri di un approccio possono semplicemente essere trasformati nei parametri dell'altro (vedi Appendice A). L'effect code implica che :

$$\sum_{A=1}^{I} \lambda^{A} = \sum_{B=1}^{J} \lambda^{B} = \sum_{C=1}^{K} \lambda^{C} = 0$$

$$\sum_{A=1}^{I} \lambda^{AB} = \sum_{B=1}^{J} \lambda^{AB} = \sum_{A=1}^{I} \lambda^{AC} = \sum_{C=1}^{K} \lambda^{AC} = \sum_{B=1}^{J} \lambda^{BC} = \sum_{C=1}^{K} \lambda^{BC} = 0$$

$$\sum_{A=1}^{I} \lambda^{ABC} = \sum_{B=1}^{J} \lambda^{ABC} = \sum_{C=1}^{K} \lambda^{ABC} = 0$$

Il dummy code implica invece che una modalità di volta in volta sia presa come baseline con la quale confrontare le altre. Prendendo ad esempio la modalità 1 per ogni variabile si ottiene:

$$\lambda^{A=1} = \lambda^{B=1} = \lambda^{C=1} = 0$$

$$\lambda^{A=1\,B=j} = \lambda^{A=iB=1} = \lambda^{A=1\,C=k} = \lambda^{A=iC=1} = \lambda^{B=1C=k} = \lambda^{B=j\,C=1} = 0$$

$$\lambda^{A=1\,B=j\,C=k} = \lambda^{A=i\,B=1\,C=k} = \lambda^{A=i\,B=j\,C=1} = 0$$

Per stimare i parametri τ (o λ) si usa il metodo della massima verosimiglianza assumendo una distribuzione multinomiale o prodotto di multinomiale. Massimizzando la funzione di verosimiglianza, non si ottengono direttamente i parametri, ma le frequenze, cioè $\hat{F}^{A=i,B=j,C=k}$, e da queste si ricavano, per via algebrica, i parametri τ e λ. I parametri sono funzione delle frequenze/probabilità. In un modello saturo (Hagenaars 1990):

$$\tau = \left(\prod_{i}\prod_{j}\prod_{k} F^{ABC}\right)^{1/IJK}$$

$$\tau^{A=i} = \frac{\left(\prod_j \prod_k F^{ABC}\right)^{1/JK}}{\tau}\,;\tau^{B=j} = \frac{\left(\prod_i \prod_k F^{ABC}\right)^{1/IK}}{\tau}\,;\tau^{C=k} = \frac{\left(\prod_i \prod_j F^{ABC}\right)^{1/IJ}}{\tau}$$

$$\tau^{A=iB=j} = \frac{\left(\prod_k F^{ABC}\right)^{1/K}}{\left(\tau\tau^{A=i}\tau^{B=j}\right)};\tau^{A=iC=k} = \frac{\left(\prod_j F^{ABC}\right)^{1/J}}{\left(\tau\tau^{A=i}\tau^{C=k}\right)};\tau^{C=kB=j} = \frac{\left(\prod_i F^{ABC}\right)^{1/I}}{\left(\tau\tau^{C=k}\tau^{B=j}\right)}$$

$$\tau^{A=iB=jC=k} = F^{ABC} \Big/ \left(\tau\tau^{A=i}\tau^{B=j}\tau^{C=k}\tau^{A=iB=j}\tau^{A=i\,C=k}\tau^{C=kB=j}\right)$$

Come già detto, si utilizza il metodo di massima verosimiglianza per stimare il modello, in quanto, sotto condizioni di regolarità, gli stimatori ML sono consistenti, asintoticamente corretti, efficienti e distribuiti asintoticamente normalmente. In un modello saturo la frequenza attesa stimata con il metodo ML è uguale a quella osservata, questo non avviene in un modello non saturo. Esistono due procedure numeriche per stimare le frequenze attese: la IPF (iterative proportional fitting) e la Newton-Raphson. Il metodo IPF è usato nei modelli gerarchici. Il suo vantaggio è che, utilizzato nel modello gerarchico, offre semplici espressioni per le statistiche sufficienti che possono essere usate per stimare i parametri del modello. Un modello gerarchico si interpreta come un modello in cui le informazioni, contenute in certe frequenze marginali e in combinazione con restrizioni messe a priori, sono sufficienti a descrivere le variazioni della tabella totale. Le marginali sono statistiche sufficienti per la stima di un particolare modello. Per esempio, in un modello ABC senza il parametro a 3 variabili, le equazioni di verosimiglianza per stimare i parametri mostrano che le frequenze attese sono quelle osservate nel campione:

$$\hat{F}^{A=i\,B=j} = F^{A=i\,B=j}, \hat{F}^{B=j\,C=k} = F^{B=j\,C=k}, \hat{F}^{A=i\,C=k} = F^{A=i\,C=k}$$

1.2 I MODELLI CAUSALI: un’evoluzione del modello log-lineare

I modelli log-lineari descritti nella parte introduttiva non spiegano relazioni di tipo causale tra le variabili. Un metodo per inserire tali relazioni in un modello log-lineare è stato introdotto per la prima volta da Goodman (1973) che lo ha chiamato “modified path model”. Si ipotizzi di avere 3 variabili tali che:

A⟶B⟶C

La variabile A è una variabile esogena, mentre C e B sono variabili endogene. Ciascuna variabile endogena può essere direttamente causalmente influenzata dalle variabili che la precedono nell’ordine causale, ma non dalle variabili che la seguono. Per esempio, B non può essere influenzato da C. Come nel modello tradizionale, anche qui gioca un ruolo importante la tavola di frequenza /di probabilità multivariata:

$$\pi^{A=i\,B=j\,C=k} = \pi^{A=i}\pi^{B=j\,|A=i}\pi^{C=k|A=i\,B=j}$$

Dopo aver determinato l’ordine gerarchico fra le variabili, si creano gli appropriati modelli log-lineari saturi/non saturi per la probabilità della variabile esogena A e per le condizionate B|A e C|AB. Come affermato da Croon et al (2009) in un modello senza variabili latenti o con variabili latenti ma senza dati clustered (cioè dataset in cui le osservazioni sono indipendenti e non raggruppate; un esempio , presente in Croon et al (2009), è un dataset composto da osservazioni su coppie genitori-figli) stimare il modello completo o stimare singolarmente i 3 sottomodelli porta agli stessi valori dei parametri; infatti i due seguenti processi di stima portano allo stesso risultato:

1) si stimano congiuntamente le probabilità condizionate $\pi^{A=i}, \pi^{B=j\,|A=i}$ e $\pi^{C=k|A=i\,B=j}$ dalla tabella di frequenza ABC;

2) si stima la probabilità $\pi^{A=i}$ dalla tabella di frequenza A, poi si stima la probabilità condizionata $\pi^{B=j\,|A=i}$ dalla tabella di frequenza AB e per ultimo si stima la probabilità condizionata $\pi^{C=k|A=i\;B=j}$ dalla tabella di frequenza ABC.

In questa capitolo si mostra che se erroneamente si stima un modello causale presupponendo un modello ABC o se si stima un modello ABC presupponendo uno causale, i parametri stimati non sono quelli reali con cui si è creato il dataset. E' proprio per fare un confronto fra i parametri stimati e quelli "reali" che si sono utilizzati dati creati esattamente, cioè senza alcuna variabilità. L'obiettivo è comprendere l'interpretazione dei parametri risultanti dalla stima conoscendo il vero processo generatore dei dati.
Le stime sono ottenute con il programma Lem (Vermunt, 1997) e per maggiori dettagli sul programma si veda l'Appendice B. I dati sono inseriti nel programma sotto forma di tabella di frequenza, e poi si è moltiplicata la probabilità per 1000 senza arrotondare (mentre le tabelle sono tutte arrotondate), ipotizzando che questa sia la numerosità campionaria. Per semplicità si è scelto di generare variabili dicotomiche, comunque risultati equivalenti si ottengono con variabili politomiche (a più categorie).

1.3 ANALISI BIVARIATA

Si inizia con la creazione del caso più semplice, cioè un dataset formato da 2 sole variabili che possono essere dipendenti o indipendenti. Se sono indipendenti si è in presenza dell'unico caso in cui si arriva allo stesso risultato, che naturalmente è esatto, sia stimando il dataset presupponendo un modello causale sia stimando il dataset presupponendo un modello AB. Se A e B sono dipendenti, stimare il dataset presupponendo un modello causale o stimare lo stesso dataset presupponendo un modello AB porta a risultati differenti. Quale delle 2 stime porti ad un'analisi corretta dipende dai dati, cioè se sono costruiti in maniera causale o meno. "Costruiti in maniera causale" significa che si presuppone di scegliere che una variabile sia esogena e che una sia endogena. Negli esempi che seguono si sceglierà A esogena e B endogena, con il seguente processo generatore dei dati

P(A=i)

P(B=j|A=i)

"Costruiti in maniera AB" significa che non si presuppone che ci sia una variabile esogena e una endogena, bensì si costruisce direttamente la congiunta. Come già detto il dataset è formato dalla tabella di frequenza e quindi per fare una stima è necessaria la congiunta di A e B (P(A=i,B=j)):

Modello causale: P(A=i,B=j) = P(A=i)P(B=j|A=i)

Modello AB: P(A=i,B=j)

1.3.1 Variabili indipendenti

Questo è il caso in cui "sbagliare il modello" con cui si stimano i dati non porta ad alcun errore, cioè i valori dei parametri sono uguali. Si creano le simulazioni non dalle probabilità ma dai parametri moltiplicativi, ipotizzando di usare l'effect code come metodo di identificazione. Si è scelto di crearle dai parametri e non dalle probabilità perché così si conoscono i reali valori dei parametri e si possono confrontare con quelli stimati. Per creare la tabella di frequenza è necessario generare la congiunta (P(A=i,B=j)) delle 2 variabili dicotomiche A e B. Per ipotesi A e B sono indipendenti. Essendo A e B variabili dicotomiche questo significa che:

$$\tau^{A=2} = (\tau^{A=1})^{-1}$$

$$\tau^{B=2} = (\tau^{B=1})^{-1}$$

$$\tau^{A=1\ B=2} = \tau^{A=2\ B=1} = (\tau^{A=1\ B=1})^{-1} = (\tau^{A=2\ B=2})^{-1}$$

I parametri scelti sono:

$$\tau^{A=1} = 0.25$$

$$\tau^{B=1} = 0.8$$
$$\tau^{A=1\ B=1} = 1$$

Da cosa si deduce che A e B sono indipendenti? La "misura" della relazione tra A e B è il parametro $\tau^{A=1\ B=1}$. Se esso è uguale ad 1 significa che le variabili A e B sono indipendenti, se diverso da 1 sono dipendenti. La tabella da cui si deriva quella di frequenza diviene:

Tabella 4: A e B indipendenti

A \ B	1	2	tot
1	0.023	0.0359	0.0589
2	0.3672	0.5739	0.9411
tot	0.3902	0.6098	1

Determinato il dataset si passa alla stima del modello.

Esistono 2 modi per stimare il modello

1)"AB"

Si ipotizza:

$$\pi^{AB} = \eta^{AB}\tau^{A}\tau^{B}\tau^{AB}$$

dove

η^{AB} rappresenta una costante di normalizzazione

τ^{A}, τ^{B} rappresentano i "one variable effects"

τ^{AB} rappresenta l'effetto interazione

Da come sono stati creati i dati, τ^{AB}, cioè "l'effetto interazione" è uguale ad 1, essendo le 2 variabili indipendenti

Tabella 5: stima del dataset "A e B indipendenti" ipotizzando modello AB

`*** LOG-LINEAR PARAMETERS ***`
 	
`* TABLE AB [or P(AB)] *`
 	
`effect lambda exp(lambda)`
`A`
`1 -1.3863 0.2500`
`2 1.3863 4.0000`
`B`
`1 -0.2231 0.8000`
`2 0.2231 1.2500`
`AB`
`1 1 0.0000 1.0000`
`1 2 0.0000 1.0000`
`2 1 0.0000 1.0000`
`2 2 0.0000 1.0000`
	`*** (CONDITIONAL) PROBABILITIES ***`
	
	`* P(AB) *`
	
	`1 1 0.0230`
	`1 2 0.0359`
	`2 1 0.3673`
	`2 2 0.5739`

Le τ sono funzione delle λ: $\tau=\exp(\lambda)$. Dalla stima risulta che $\hat{\tau}^{AB}$è effettivamente uguale a 1, come quello "reale", lo stesso vale per i parametri $\hat{\tau}^{A=1} = 0.25$ e $\hat{\tau}^{B=1} = 0.8$.[1] La costante di normalizzazione è $\hat{\eta}^{AB}$=0.1148. Essa serve affinché la sommatoria delle probabilità della congiunta sia uguale ad 1 e quindi si parte dalla seguente identità per derivarla:

$$\sum_B \sum_A P(A = i, B = j) \equiv 1$$

Ricordando che P(A=i,B=j)= $\eta^{AB}\tau^{B=j}\tau^{A=i}\tau^{A=i\,B=j}$, la sommatoria diviene

$$\sum_B \sum_A \eta^{AB}\tau^{B=j}\tau^{A=i}\tau^{A=i\,B=j} = 1$$

[1] Si analizzano dati "non casuali" e quindi nessun tipo di test può essere fatto. Se i dati sono casuali esiste un test di indipendenza delle 2 variabili: che è il test sul parametro λ^{AB} , la trasformazione logaritmica di τ^{AB}. Che ipotesi deve essere fatta sul parametro λ^{AB} per studiare l'indipendenza? Se si sa che $\tau^{AB} = 1$ significa che A e B sono indipendenti, allora per studiare l'indipendenza si pone come ipotesi nulla del test che λ^{AB} sia uguale a 0. Esistono due tipi di test di indipendenza, che nel caso dicotomico si equivalgono: lo z-test e il test di Wald. Entrambi i test sono spiegati in maniera più dettagliata nel terzo capitolo , dove è spiegato come studiare la significatività dei parametri.

Raccogliendo a fattor comune risulta

$$\eta^{AB}(\tau^{B=1}\tau^{A=1}\tau^{A=1,B=1} + \tau^{B=2}\tau^{A=1}\tau^{A=1,B=2} + \tau^{B=1}\tau^{A=2}\tau^{A=2,B=1} + \tau^{B=2}\tau^{A=2}\tau^{A=2,B=2})=1$$

e quindi

$$\eta^{AB} = \frac{1}{\tau^{A=1}\tau^{B=1}\tau^{A=1\,B=1} + \tau^{A=1}\,\tau^{B=2}\tau^{A=1\,B=2} + \tau^{A=2}\tau^{B=1}\tau^{A=2\,B=1} + \tau^{A=2}\tau^{B=2}\tau^{A=2\,B=2}}$$

e nel caso specifico

$$\widehat{\eta}^{AB} = \frac{1}{0.25\;0.8 + 0.25\;1.25 + 4\;0.8 + 4\;1.25} = \frac{1}{8.7125} = 0.1148$$

2)”A e B|A”

Si ipotizza una gerarchia tra variabili (modello causale):

$$\pi^{A} = \eta^{A}\tau^{A}$$

$$\pi^{AB} = \eta^{AB}\tau^{A}\tau^{B}\tau^{AB} \Rightarrow \pi^{B|A} = \eta^{B|A}\tau^{B}\tau^{AB}$$

Tabella 6: stima dataset “A e B indipendenti” ipotizzando modello causale

<pre>*** LOG-LINEAR PARAMETERS ***
 	
* TABLE A [or P(A)] *
 	
  effect   lambda exp(lambda)
  A
   1   -1.3863   0.2500
   2   1.3863   4.0000
 	
* TABLE AB [or P(B|A)] *
 	
  effect   lambda exp(lambda)
  B
   1   -0.2231   0.8000
   2   0.2231   1.2500
  AB
   1 1   0.0000   1.0000
   1 2   0.0000   1.0000
   2 1   0.0000   1.0000
   2 2   0.0000   1.0000</pre>
	<pre>*** (CONDITIONAL) PROBABILITIES ***
	
	* P(A) *
	
	  1 0.0588
	  2 0.9412
	
	* P(B|A) *
	
	  1 | 1 0.3902
	  2 | 1 0.6098
	  1 | 2 0.3902
	  2 | 2 0.6098</pre>

Dalla stima risulta che $\hat{\tau}^{AB}$è effettivamente uguale ad 1, come quello “reale” e ciò anche in questo modello, che è gerarchico, vuol dire indipendenza tra le 2 variabili, allo stesso modo si ha $\widehat{\tau}^{A=1} = 0.25$ e $\widehat{\tau}^{B=1} = 0.8$. Le costanti di normalizzazione sono: $\widehat{\eta}^{B|A}$=0.4878, $\widehat{\eta}^{BA}$=0.1147 e $\widehat{\eta}^{A}$=0.2353. La costante $\widehat{\eta}^{A}$ si usa per assicurare che la somma delle marginali di A dia 1 come risultato; $\widehat{\eta}^{B|A}$ si usa per assicurare che la somma delle condizionate di B dato A dia 1 come risultato e $\widehat{\eta}^{BA}$ rappresenta il prodotto tra le $2(\widehat{\eta}^{A}\ \widehat{\eta}^{B|A})$. Di seguito sono elencate procedure generali per ricavare $\widehat{\eta}^{A}, \widehat{\eta}^{B|A}$ e $\widehat{\eta}^{BA}$.

- **Per ricavare $\widehat{\eta}^{A}$** si parte dall'identità

$\sum_A P(A = i) \equiv 1$

Nei modelli log-lineari P(A=i) è uguale a $\eta^{A}\tau^{A}$ e quindi

$$\sum_A \eta^{A}\tau^{A} = 1$$

Si raccoglie a fattor comune

$$\eta^{A}(\tau^{A=1} + \tau^{A=2}) = 1$$

e si ricava così η^{A}

$$\eta^{A} = \frac{1}{(\tau^{A=1} + \tau^{A=2})} \qquad (3)$$

Nel caso specifico

$$\widehat{\eta}^{A} = \frac{1}{0.25 + 4} = \frac{1}{4.25} = 0.2353$$

Si può esplicitare $\eta^{A=1}$ come funzione di un solo parametro e precisamente $\tau^{A=1}$ grazie alla parametrizzazione effect code. Ricordando che

$$\tau^{A=2} = \frac{1}{\tau^{A=1}}$$

Questa si sostituisce nella (3), che diviene

$$\eta^{A} = \frac{1}{(\tau^{A=1} + \tau^{A=2})} = \frac{1}{\tau^{A=1} + \frac{1}{\tau^{A=1}}} = \frac{\tau^{A=1}}{(\tau^{A=1})^2 + 1}$$

Nel caso specifico

$$\widehat{\eta}^{A} = \frac{0.25}{0.25^2 + 1} = \frac{0.25}{1.0625} = 0.2353$$

che ha lo stesso risultato

- **Per ricavare $\widehat{\eta}^{B|A=i}$** si parte dall'identità

$\sum_B P(B = j|A = i) \equiv 1 \qquad (4)$

Si è specificato A=i nella $\widehat{\eta}^{B|A=i}$ perché nelle probabilità condizionate esistono tante $\widehat{\eta}^{B|A=i}$ quante sono le categorie di A ; poiché A è dicotomica si hanno 2 valori di $\widehat{\eta}^{B|A}$.

Nel caso di indipendenza tra A e B i 2 valori delle $\widehat{\eta}^{B|A}$ coincidono. Si nota che in $\widehat{\eta}^{B|A=i}$ non è stato posto B=j perché $\widehat{\eta}^{B=j|A=i} = \widehat{\eta}^{B=h|A=i}$ per ogni h≠j.

Ricordando che

P(B=J|A=i)= $\eta^{B|A=i}\tau^{B=j}\tau^{A=iB=j}$

si sostituisce P(B=j|A=i) nella identità (4), che diviene

$$\sum_{B} \eta^{B|A=i}\tau^{B=j}\tau^{A=iB=j} = 1$$

Raccogliendo a fattor comune

$\eta^{B|A=i}\left(\tau^{B=1}\tau^{A=iB=1} + \tau^{B=2}\tau^{A=iB=2}\right)$=1

si ricava $\eta^{B|A=i}$:

$$\eta^{B|A=i} = \frac{1}{\tau^{B=1}\tau^{A=iB=1}+\tau^{B=2}\tau^{A=iB=2}} \qquad (5)$$

Nel caso specifico

$$\eta^{B|A=2} = \frac{1}{0.8 + 1.25} = \frac{1}{2.05} = 0.4878$$

$$\eta^{B|A=1} = \frac{1}{0.8 + 1.25} = \frac{1}{2.05} = 0.4878$$

E' possibile arrivare ad esplicitare $\eta^{B|A=1}$ come funzione di 2 soli parametri $\tau^{B=1}$ e $\tau^{A=1\,B=1}$ grazie alla parametrizzazione effect code. Ricordando che

$\tau^{B=2} = \frac{1}{\tau^{B=1}}$ e $\tau^{A=1\,B=1} = \tau^{A=2\,B=2} = \frac{1}{\tau^{A=2\,B=1}} = \frac{1}{\tau^{A=1\,B=2}}$

la (5) diviene

$$\eta^{B|A=i} = \frac{1}{\tau^{B=1}\tau^{A=iB=1}+\tau^{B=2}\tau^{A=iB=2}} = \frac{1}{\tau^{B=1}\tau^{A=iB=1}+\frac{1}{\tau^{B=1}\tau^{A=iB=1}}}$$

$$=> \eta^{B|A=i} = \frac{\tau^{B=1}\tau^{A=iB=1}}{\left(\tau^{B=1}\tau^{A=iB=1}\right)^2+1}$$

Nel caso specifico

$$\eta^{B|A=1} = \frac{0.8}{0.8^2 + 1} = \frac{0.8}{1.64} = 0.4878$$

- **Ora si può calcolare $\widehat{\eta}^{BA}$:**

$\widehat{\eta}^{BA} = \eta^{B|A=i}\eta^{A}$=0.4878x0.2353=0.1147

1.3.2 Variabili dipendenti non causalmente

Stimando un dataset creato in maniera causale con il metodo AB e un dataset creato con il metodo AB con il metodo causale, i parametri stimati risultano essere differenti da quelli con cui si sono creati i dati. I parametri $\tau^{A=1}$e $\tau^{B=1}$ rimangono gli stessi del modello con A e B indipendenti, si cambia solo il valore di $\tau^{A=1\,B=1}$, proprio per rendere A e B dipendenti. I parametri scelti , che continuano ad essere identificati con il modello effect code, sono quindi:

$$\tau^{A=1} = 0.25$$

$$\tau^{B=1} = 0.8$$
$$\tau^{A=1,B=1} = 0.5$$

L'effetto interazione è diverso da 0, A e B sono dipendenti, ma non si ipotizza un ordine tra loro, cioè si crea direttamente la distribuzione congiunta AB, costruendo η^{AB} in modo tale che la somma delle probabilità congiunte sia pari all'unità, quindi utilizzando la formula già usata per il modello AB:

$$\eta^{AB} = \frac{1}{0.25\ 0.8\ 0.5 + 0.25\ 1.25\ 2 + 4\ 0.8\ 2 + 4\ 1.25\ 0.5} = \frac{1}{9.625} = 0.1039$$

La probabilità congiunta di A e B è così calcolata

P(A=i,B=j)= $\eta^{AB}\tau^{A=i}\tau^{B=j}\tau^{A=i\,B=j}$=0.1039 $\tau^{A=i}\tau^{B=j}\tau^{A=i\,B=j}$

e quindi per esempio

P(A=1,B=1)=0.1039X0.25X0.8X0.5=0.0104

Tabella 7: A e B dipendenti ma non in maniera causale

B \ A	1	2	tot
1	0.0104	0.0650	0.0754
2	0.6649	0.2597	0.9246
tot	0.6753	0.3247	1

Dalla tabella 8 si vede che esiste un unico modo per stimare il modello dalla stessa congiunta:

Tabella 8: stima dataset "A e B dipendenti ma non in maniera causale" ipotizzando due modelli differenti

Stima ipotizzando modello AB o modello log-lineare tradizionale	Stima ipotizzando modello causale

`*** LOG-LINEAR PARAMETERS ***`
 	
`* TABLE AB [or P(AB)] *`
 	
`effect          lambda   exp(lambda)`
`A`
` 1              -1.3863    0.2500`
` 2               1.3863    4.0000`
`B`
` 1              -0.2231    0.8000`
` 2               0.2231    1.2500`
`AB`
` 1 1            -0.6931    0.5000`
` 1 2             0.6931    2.0000`
` 2 1             0.6931    2.0000`
` 2 2            -0.6931    0.5000`
	`*** LOG-LINEAR PARAMETERS ***`
	
	`* TABLE A [or P(A)] *`
	
	`effect          lambda   exp(lambda)`
	`A`
	` 1              -1.2538    0.2854`
	` 2               1.2538    3.5037`
	`* TABLE AB [or P(B|A)] *`
	`effect          lambda   exp(lambda)`
	`B`
	` 1              -0.2231    0.8000`
	` 2               0.2231    1.2500`
	`AB`
	` 1 1            -0.6931    0.5000`
	` 1 2             0.6931    2.0000`
	` 2 1             0.6931    2.0000`
	` 2 2            -0.6931    0.5000`

Il secondo metodo (colonna 2 della tabella) non è più utilizzabile, perché presuppone una gerarchia/ causalità tra le variabili; infatti esso ipotizza che A sia esogena e B endogena. Se si stima il modello con questo metodo, B si stima correttamente, a differenza di A che si stima in maniera erronea (i risultati sono riportati nella colonna 2 della tabella 8): rimangono comunque identificati correttamente i parametri $\tau^{B=j} = 0.8 \; e \; \tau^{A=i\,B=j} = 0.5$. Nel modello causale i parametri sono suddivisi in P(A) e P(B|A), mentre nel modello AB i parametri sono tutti insieme a causa della differenza teorica dei due modelli. Nel modello AB i parametri $\tau^{A=i}\tau^{B=j}\tau^{A=i\,B=j}$ sono tutti ricavati dalla stessa tabella di frequenza AB. Nel modello causale $\tau^{A=i}$ è ricavato dalla tabella di frequenza A, mentre $\tau^{B=j} \; e \; \tau^{A=i\,B=j}$ sono ricavati dalla tabella B|A:

- <u>modello AB o loglineare tradizionale</u>, i parametri si ricavano da $F^{A=i\,B=j}$, cioè

$$\frac{F^{A=i\,B=j}}{N} = P(A = i\; B = j) = \eta^{AB}\tau^{A=i}\tau^{B=j}\tau^{A=i\,B=j}$$

- <u>modello causale,</u> il parametro $\tau^{A=i}$ si ricava da $F^{A=i}$ e i parametri $\tau^{B=j} \; e \; \tau^{A=i\,B=j}$ si ricavano da $F^{B=j|A=i}$,cioè

$$\frac{F^{A=i}}{N} = P(\text{A=i}) = \eta^{A}\tau^{A=i}$$

$$\frac{F^{B=j|A=i}}{N} = P(B = j|A = i) = \eta^{B|A=i}\tau^{B=j}\tau^{A=i\,B=j} \qquad (6)$$

La conclusione è che nei dataset in cui le variabili sono dipendenti non in maniera causale , se stimati ipotizzando un modello causale, le stime non sono identificate correttamente.

1.3.3 Variabili dipendenti causalmente

Ora si analizza un dataset con le variabili dipendenti in maniera causale. Se è stimato ipotizzando un modello non causale, ci si chiede se le stime siano uguali. Con i medesimi valori dei parametri utilizzati per costruire il dataset della tabella 7, si crea la distribuzione di A e poi quella di B|A:

Tabella 9: marginale di A e condizionata di B

A		B|A	
1	0.0588	1|1	0.1379
2	0.9412	2|1	0.8621
		1|2	0.7191
		2|2	0.2809
η^{A}=0.2353		$\eta^{B\|A=1} = 0.3448$ $\eta^{B\|A=2} = 0.4494$	

Per esempio:

P(A=1)= $\eta^{A}\tau^{A=1} = 0.2353x0.25 = 0.0588$

P(B=1|A=1)= $\eta^{B|A=1}\tau^{B=1}\tau^{A=1\ B=1} = 0.3448x0.8x0.5 = 0.1379$

Come sono costruiti gli η? Nel modello precedente, cioè quello delle variabili con dipendenza non causale, l'η è una costante di normalizzazione che assicura che la somma di tutte le probabilità sia uguale ad 1. Quando si introducono le condizionate, le η servono, invece, per garantire che:

$$P(B = 1|A) + P(B = 2|A) = 1$$

e quindi ricordando i passaggi precedenti, cioè

$$\eta^{B|A=i}\left(\tau^{B=1}\tau^{A=i\ B=1} + \tau^{B=2}\tau^{A=i\ B=2}\right)=1$$

risulta

$$\eta^{B|A=i} = \frac{1}{\tau^{B=1}\tau^{A=i\ B=1} + \tau^{B=2}\tau^{A=i\ B=2}}$$

Nel caso specifico:

$$\eta^{B|A=1} = \frac{1}{0.8\ 0.5 + 1.25\ \ 2} = \frac{1}{2.9} = 3.448$$

$$\eta^{B|A=2} = \frac{1}{0.8\ \ 2 + 1.25\ \ 0.5} = \frac{1}{2.225} = 0.4494$$

Se si utilizza la formula che tiene in considerazione la parametrizzazione effect code

$$\eta^{B|A=i} = \frac{\tau^{B=1}\tau^{A=i\ B=1}}{\left(\tau^{B=1}\tau^{A=i\ B=1}\right)^2+1}$$

Risulta

$$\eta^{B|A=1} = \frac{0.8\ 0.5}{(0.8\ 0.5)^2 + 1} = \frac{0.4}{1.16} = 0.3448$$

$$\eta^{B|A=2} = \frac{0.8\ 2}{(0.8\ 2)^2 + 1} = \frac{1.6}{3.56} = 0.4494$$

Ora che si sono ricavate P(A=i) e P(B=j|A=i) (valori della tabella 9), si calcola la distribuzione congiunta P(A=i,B=j), ricordando che

P(A=i,B=j)= P(A=i) P(B=j|A=i)

Per esempio

P(A=1,B=1)=0.0588X0.1379=0.0081

Le probabilità congiunte così ottenute sono descritte nella tabella 10:

Tabella 10: A e B dipendenti causalmente

B / A	1	2	Tot
1	0.0081	0.0507	0.0588
2	0.6768	0.2644	0.9412
tot	0.6849	0.3151	1

Questa volta il metodo corretto per stimare i dati è il causale (colonna 1 della tabella 11)

Tabella 11: stima del dataset "A e B dipendenti causalmente" ipotizzando due modelli differenti

<table>
<tr><td>Stima ipotizzando modello causale</td><td>Stima ipotizzando modello AB o modello log-lineare tradizionale</td></tr>
<tr><td><pre>*** LOG-LINEAR PARAMETERS ***

* TABLE A [or P(A)] *

  effect   lambda exp(lambda)
  A
   1   -1.3863   0.2500
   2   1.3863   4.0000

* TABLE AB [or P(B|A)] *

  effect   lambda exp(lambda)
  B
   1   -0.2231   0.8000
   2   0.2231   1.2500
  AB
   1 1   -0.6931   0.5000
   1 2   0.6931   2.0000
   2 1   0.6931   2.0000
   2 2   -0.6931   0.5000</pre></td><td><pre>*** LOG-LINEAR PARAMETERS ***

* TABLE AB [or P(AB)] *

  effect   lambda exp(lamba)
  A
   1   -1.5188   0.2190
   2   1.5188   4.5666
  B
   1   -0.2231   0.8000
   2   0.2231   1.2500
  AB
   1 1   -0.6931   0.5000
   1 2   0.6931   2.0000
   2 1   0.6931   2.0000
   2 2   -0.6931   0.5000</pre></td></tr>
</table>

Il secondo metodo (colonna 2 della tabella) non è più utilizzabile, perché non presuppone una gerarchia/ causalità tra le variabili. Se si stima il modello con questo metodo, B si stima esattamente, a differenza di A che si stima in maniera erronea (i risultati sono riportati nella colonna 2 della tabella 11): rimangono comunque identificati correttamente i parametri $\tau^{B=j} = 0.8$ e $\tau^{A=i\,B=j} = 0.5$.

1.3.4 Conclusioni

I tre modelli visti nelle 3 sezioni precedenti sono sempre costruiti con $\tau^{A} = 0.25$ e τ^{B}=0.8. Nel passaggio dal primo modello (indipendenza) agli altri 2 modelli (dipendenza non causale, dipendenza causale) si modifica τ^{AB}. La differenza tra i due modelli con dipendenza non sta nei parametri che rimangono uguali ma nella costruzione dei dati. Con A e B indipendenti, la stima è la medesima. Con A e B dipendenti, è necessario , stimando , capire il vero processo che ha generato i dati per poter avere le stime uguali ai valori utilizzati nella creazione del dataset; infatti per stimare i dati con dipendenza causale si presuppone un modello causale, per stimare i dati con dipendenza non causale si presuppone un modello AB. Il problema principale è nella definizione delle costanti di normalizzazione e nella marginale della variabile esogena, mentre il parametro di interesse sulla relazione fra A e B rimane uguale nelle 2 procedure.

1.4 MODELLO CAUSALE MULTIVARIATO:i sub modelli

Dopo aver stimato i modelli a due variabili dicotomiche (A e B), si prosegue con quelli con tre variabili dicotomiche (A, B e C). Si utilizza come metodo di identificazione l'effect code. Dopo aver creato un modello ABC e a averlo stimato sia ipotizzando un modello causale sia un modello log-lineare "tradizionale", si studia un modello ricorsivo completo. Da quest'ultimo si analizza il significato dei vari parametri.

1.4.1 Modello ABC e modello ricorsivo completo

Inizialmente si creano 3 variabili dipendenti, calcolando direttamente la loro distribuzione congiunta dai seguenti parametri:

$$\tau^{A} = 0.25$$

$$\tau^{B} = 0.8$$
$$\tau^{C} = 0.2$$

$$\tau^{AB} = 0.5$$

$$\tau^{AC} = 1.6$$

$$\tau^{CB} = 3.2$$

$$\tau^{CAB} = 1$$

I parametri τ^{A}, τ^{B} e τ^{AB} rimangono gli stessi del modello con A e B dipendenti (tabella 7), si aggiungono solo i parametri riguardanti la nuova variabile.

La probabilità congiunta è rappresentata nella tabella 12 che non contiene dati sparsi in quanto tutte le celle hanno almeno una osservazione (Heinen 1996).

Tabella 12: A, B e C dipendenti non causalmente

cab	
111	0.0011
112	0.0007
121	0.0287
122	0.0011
211	0.0012
212	0.0701
221	0.1794
222	0.7177
tot	1
$\eta^{ABC} = 0.0112$	

La η^{ABC} viene calcolata dalla seguente formula:

$$\eta^{ABC}(\sum_C \sum_B \sum_A \tau^B \tau^A \tau^C \tau^{AB} \tau^{AC} \tau^{CB})=1$$

Questo modello deve essere stimato congiuntamente, cioè stimando insieme la probabilità congiunta ABC, poiché non c'è una gerarchia nella costruzione delle variabili. La stima corretta è rappresentata nella colonna 1 della tabella 13.

Tabella 13:stima del dataset "A, B e C dipendenti non causalmente " ipotizzando due modelli differenti

<table>
<tr><th>Stima ipotizzando modello ABC o modello log-lineare tradizionale</th><th>Stima ipotizzando modello causale</th></tr>
<tr><td><pre>
*** LOG-LINEAR PARAMETERS ***

* TABLE CAB [or P(CAB)] *

  effect   lambda exp(lambda)
  A
   1   -1.3863   0.2500
   2   1.3863   4.0000
  B
   1   -0.2231   0.8000
   2   0.2231   1.2500
  C
   1   -1.6094   0.2000
   2   1.6094   5.0000
  AB
   1 1   -0.6931   0.5000
   1 2   0.6931   2.0000
   2 1   0.6931   2.0000
   2 2   -0.6931   0.5000
  CA
   1 1   0.4700   1.6000
   1 2   -0.4700   0.6250
   2 1   -0.4700   0.6250
   2 2   0.4700   1.6000
  CB
   1 1   1.1632   3.2000
   1 2   -1.1632   0.3125
   2 1   -1.1632   0.3125
   2 2   1.1632   3.2000

  CAB
   1 1 1   0.0000   1.0000
   1 1 2   0.0000   1.0000
   1 2 1   0.0000   1.0000
   1 2 2   0.0000   1.0000
   2 1 1   0.0000   1.0000
   2 1 2   0.0000   1.0000
   2 2 1   0.0000   1.0000
   2 2 2   0.0000   1.0000
</pre></td><td><pre>
*** LOG-LINEAR PARAMETERS ***

* TABLE A [or P(A)] *

  effect   lambda exp(lambda)
  A
   1   -1.2705   0.2807
   2   1.2705   3.5626

* TABLE AB [or P(B|A)] *

  effect   lambda exp(lambda)
  B
   1   -1.1728   0.3095
   2   1.1728   3.2309
  AB
   1 1   -0.5531   0.5752
   1 2   0.5531   1.7386
   2 1   0.5531   1.7386
   2 2   -0.5531   0.5752

* TABLE CAB [or P(C|AB)] *

  effect   lambda exp(lambda)
  C
   1   -1.6094   0.2000
   2   1.6094   5.0000
  CA
   1 1   0.4700   1.6000
   1 2   -0.4700   0.6250
   2 1   -0.4700   0.6250
   2 2   0.4700   1.6000
  CB
   1 1   1.1632   3.2000
   1 2   -1.1632   0.3125
   2 1   -1.1632   0.3125
   2 2   1.1632   3.2000
  CAB
   1 1 1   0.0000   1.0000
   1 1 2   0.0000   1.0000
   1 2 1   0.0000   1.0000
   1 2 2   0.0000   1.0000
   2 1 1   0.0000   1.0000
   2 1 2   0.0000   1.0000
   2 2 1   0.0000   1.0000
   2 2 2   0.0000   1.0000
</pre></td></tr>
</table>

Dalla stima risulta che $\hat{\tau}^{CAB}$è effettivamente uguale ad 1, come quello “reale” in entrambe le procedure, come risulta anche per $\hat{\tau}^{C=1A=1} = 1.6$, $\hat{\tau}^{C=1B=1}$=3.2 e $\hat{\tau}^{C=1} =$ 0.2.
Se si ipotizza un ordine causale, cioè ad esempio che A è indipendente, che B dipende da A e che C dipende da A e B, questo deve essere inserito nella procedura di stima. La differenza tra un modello causale e un modello CAB è che la relazione tra le variabili nel primo ha una direzione, nel secondo no, cioè:
modello causale: A⟶B →C e A→C

modello ABC: A⟷B⟷C e A⟷C

Il modello, che tradizionalmente viene usato per le variabili continue nell’analisi SEM e nelle analisi di regressione, è quello causale. Ora si crea un dataset con relazioni causali. I parametri moltiplicativi sono gli stessi di prima, ma è la relazione tra le variabili ad essere diversa.

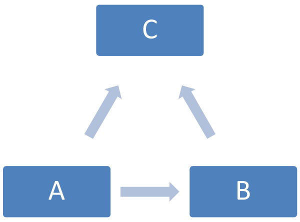


Figura 1: modello ricorsivo completo

Le distribuzioni di A e B rimangono le stesse della tabella 9, quella di C diviene

Tabella 14: condizionata di C e congiunta CAB con A, B e C dipendenti causalmente

c|ab		cab	
1|11	0.5119	111	0.0042
2|11	0.4881	211	0.0040
1|12	0.0099	112	0.0005
2|12	0.9901	212	0.0502
1|21	0.1379	121	0.0933
2|21	0.8621	221	0.5834
1|22	0.0015	122	0.0004
2|22	0.9985	222	0.2640
		tot	1
$\eta^{C\|A=1B=1} = 0.4999$ $\eta^{C\|A=1\,B=2} = 0.0990$ $\eta^{C\|A=2B=1} = 0.3448$ $\eta^{C\|A=2\,B=2} = 0.0390$			

Quindi le probabilità sono state così costruite:

$$\pi_a^A = \eta^A \tau^A$$

$$\pi_{b\,a}^{B|A} = \eta^{B|A} \tau^B \tau^{AB}$$

$$\pi_{c\;ab}^{C|AB} = \eta^{C|AB} \tau^C \tau^{CA} \tau^{CB}$$

e con questa sequenza il programma Lem le stima

Tabella 15 :stima modello "A, B e C dipendenti causalmente" ipotizzando modello causale e modello ABC

<table>
<tr><td>Stima ipotizzando modello causale</td><td>Stima ipotizzando modello ABC o modello log-lineare tradizionale</td></tr>
<tr><td>*** LOG-LINEAR PARAMETERS ***

* TABLE A [or P(A)] *

effect lambda exp(lambda)
A
1 -1.3863 0.2500
2 1.3863 4.0000

* TABLE AB [or P(B|A)] *

effect lambda exp(lambda)
B
1 -0.2231 0.8000
2 0.2231 1.2500
AB
1 1 -0.6931 0.5000
1 2 0.6931 2.0000
2 1 0.6931 2.0000
2 2 -0.6931 0.5000

* TABLE CAB [or P(C|AB)] *

effect lambda exp(lambda)
C
1 -1.6094 0.2000
2 1.6094 5.0000
CA
1 1 0.4700 1.6000
1 2 -0.4700 0.6250
2 1 -0.4700 0.6250
2 2 0.4700 1.6000
CB
1 1 1.1632 3.2000
1 2 -1.1632 0.3125
2 1 -1.1632 0.3125
2 2 1.1632 3.2000
CAB
1 1 1 0.0000 1.0000
1 1 2 0.0000 1.0000
1 2 1 0.0000 1.0000
1 2 2 0.0000 1.0000
2 1 1 0.0000 1.0000
2 1 2 0.0000 1.0000
2 2 1 0.0000 1.0000
2 2 2 0.0000 1.0000</td><td>* TABLE CAB [or P(CAB)] *

effect lambda exp(lambda)
C
1 -1.6094 0.2000
2 1.6094 5.0000
A
1 -1.1931 0.3033
2 1.1931 3.2971
B
1 0.7265 2.0678
2 -0.7265 0.4836
CA
1 1 0.4700 1.6000
1 2 -0.4700 0.6250
2 1 -0.4700 0.6250
2 2 0.4700 1.6000
CB
1 1 1.1632 3.2000
1 2 -1.1632 0.3125
2 1 -1.1632 0.3125
2 2 1.1632 3.2000
AB
1 1 -0.8332 0.4346
1 2 0.8332 2.3007
2 1 0.8332 2.3007
2 2 -0.8332 0.4346
CAB
1 1 1 0.0000 1.0000
1 1 2 0.0000 1.0000
1 2 1 0.0000 1.0000
1 2 2 0.0000 1.0000
2 1 1 0.0000 1.0000
2 1 2 0.0000 1.0000
2 2 1 0.0000 1.0000
2 2 2 0.0000 1.0000</td></tr>
</table>

```
*** (CONDITIONAL) PROBABILITIES ***

* P(A) *

  1              0.0588
  2              0.9412

* P(B|A) *

  1 | 1          0.1379
  2 | 1          0.8621
  1 | 2          0.7191
  2 | 2          0.2809

* P(C|AB) *

  1 | 1 1        0.5119
  1 | 1 2        0.0099
  1 | 2 1        0.1379
  1 | 2 2        0.0015
  2 | 1 1        0.4881
  2 | 1 2        0.9901
  2 | 2 1        0.8621
  2 | 2 2        0.9985
```

```
*** (CONDITIONAL) PROBABILITIES ***

* P(CAB) *

  1 1 1          0.0042
  1 1 2          0.0005
  1 2 1          0.0933
  1 2 2          0.0004
  2 1 1          0.0040
  2 1 2          0.0502
  2 2 1          0.5834
  2 2 2          0.2640
```

Un modello, come quello causale appena analizzato, ha l'interazione ABC nulla, ma tutte le interazioni di secondo livello e gli effetti "one variable" non nulle. Nei modelli di questa sezione non tutte le interazioni di secondo livello sono non nulle, cioè si hanno dei sotto-modelli. Il modello completo di partenza è quello del dataset della tabella 14 (la cui stima è rappresentata nella tabella 15)

1.4.2 A non influisce su C: presenza del solo effetto indiretto

Si creano 3 variabili ABC tali che A influisce su B che a sua volta influisce su C. Questa costruzione comporta che A non influisce direttamente su C, ma solo indirettamente attraverso B.

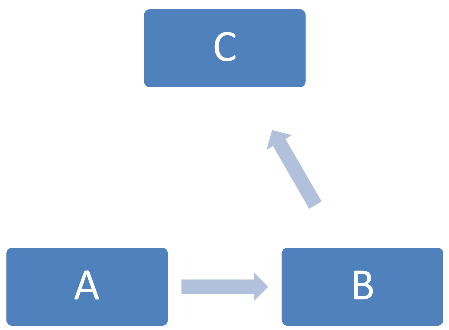


Figura 2: modello: presenza solo effetto indiretto

I parametri moltiplicativi usati per la costruzione delle variabili sono:

$$\tau^{A} = 0.25$$

$$\tau^{B} = 0.8$$

$$\tau^{C} = 0.2$$
$$\tau^{AB} = 0.5$$

$$\tau^{AC} = 1$$

$$\tau^{CB} = 3.2$$

$$\tau^{CAB} = 1$$

Dato a τ^{AC} il valore 1 , si ha l'assenza dell'effetto diretto di A su C. La creazione dei dati avviene come nel modello completo, cioè prima si calcola la probabilità di A, poi la condizionata B|A e successivamente la condizionata C|B. Proprio perché si opera a stadi, la probabilità di A, la condizionata B|A e quindi la congiunta AB rimangono le stesse del modello completo, mentre cambia la condizionata C|B e quindi la congiunta CAB:

Tabella 16: condizionata di B , di C e congiunta CAB, sub-modello "con solo effetto indiretto"

b|a		c|ab	
1|1	0.1379	1|11	0.2906
2|1	0.8621	2|11	0.7094
1|2	0.7191	1|12	0.0039
2|2	0.2809	2|12	0.9961
		1|21	0.2906
		2|21	0.7094
		1|22	0.0039
		2|22	0.9961
$\eta^{B\|A=1}$=0.3448		$\eta^{C\|A=1\ B=1} = \eta^{C\|A=2\ B=1} = 0.4540$	
$\eta^{B\|A=2}$ =0.4494		$\eta^{C\|A=1\ B=2} = \eta^{C\|A=2\ B=2} = 0.0623$	
cab			
111	0.0024		
112	0.0002		
121	0.1967		
122	0.0010		
211	0.0058		
212	0.0505		
221	0.4801		
222	0.2633		

Il modello stimato è

Tabella 17: stima modello "con solo effetto indiretto"

*** LOG-LINEAR PARAMETERS ***	*** (CONDITIONAL) PROBABILITIES ***

<pre>* TABLE A [or P(A)] *
 	
  effect   lambda exp(lambda)
  A
   1   -1.3863   0.2500
   2   1.3863   4.0000
 	
* TABLE AB [or P(B|A)] *
 	
  effect   lambda exp(lambda)
  B
   1   -0.2231   0.8000
   2   0.2231   1.2300
  AB
   1 1   -0.6931   0.5000
   1 2   0.6931   2.0000
   2 1   0.6931   2.0000
   2 2   -0.6931   0.5000
 	
* TABLE CAB [or P(C|AB)] *
 	
  effect   lambda exp(lambda)
  C
   1   -1.6094 0.2000
   2   1.6094 5.0000
  CA
   1 1   0.0000 1.0000
   1 2   -0.0000 1.0000
   2 1   -0.0000 1.0000
   2 2   0.0000 1.0000
  CB
   1 1   1.1632 3.2000
   1 2   -1.1632 0.3125
   2 1   -1.1632 0.3125
   2 2   1.1632 3.2000
  CAB
   1 1 1   0.0000 1.0000
   1 1 2   0.0000 1.0000
   1 2 1   0.0000 1.0000
   1 2 2   0.0000 1.0000
   2 1 1   0.0000 1.0000
   2 1 2   0.0000 1.0000
   2 2 1   0.0000 1.0000
   2 2 2   0.0000 1.0000</pre>
	<pre>* P(A) *
	
	  1 0.0588
	  2 0.9412
	
	* P(B|A) *
	
	  1 | 1 0.1379
	  2 | 1 0.8621
	  1 | 2 0.7191
	  2 | 2 0.2809
	
	* P(C|AB) *
	
	  1 | 1 1 0.2906
	  1 | 1 2 0.0039
	  1 | 2 1 0.2906
	  1 | 2 2 0.0039
	  2 | 1 1 0.7094
	  2 | 1 2 0.9961
	  2 | 2 1 0.7094
	  2 | 2 2 0.9961</pre>

Nella tabella 17 (colonna 1) si osservano i parametri influenti nella condizionata P(C| A, B) nella “table CAB”, quelli influenti nella condizionata P(B|A) nella “table AB” e quelli influenti nella marginale di P(A) nella “table A”. Il parametro, la cui assenza nel processo generatore dei dati determina il sotto modello, è τ^{CA} che risulta identificato uguale ad 1. L’effetto di A su B (τ^{AB}) è 0.5, quello di B su C (τ^{CB}) è 3.2. L’interazione tra le variabili A,B e C, nulla nel processo generatore dei dati, è nulla ($\tau^{CAB} = 1 => \lambda^{CAB} = 0$).

1.4.3 A non influisce su B: primo caso di presenza del solo effetto diretto

Ora si crea un modello in cui A e B influiscono su C, ma A non influisce su B (equivalente ad un modello di regressione multipla). Si continuano a generare modelli gerarchici perciò si creano prima le probabilità di A e B e poi la condizionata C|AB. Data la condizionata, si passa alla congiunta CAB semplicemente moltiplicando la condizionata per le probabilità di A e di B.

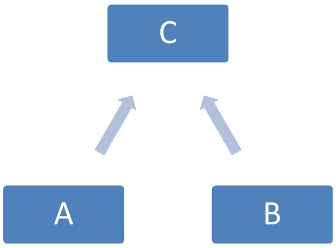


Figura 3: primo caso di presenza del solo effetto diretto

I parametri moltiplicativi utilizzati sono:

$$\tau^{A} = 0.25$$

$$\tau^{B} = 0.8$$

$$\tau^{C} = 0.2$$
$$\tau^{AB} = 1$$

$$\tau^{AC} = 1.2$$

$$\tau^{CB} = 3.2$$

$$\tau^{CAB} = 1$$

Tabella 18: condizionata di B , di C e congiunta CAB, sub-modello “primo caso con solo effetto diretto”

b|a		c|ab		cab	
1|1	0.3902	1|11	0.3710	111	0.0086
2|1	0.6098	2|11	0.6290	112	0.0002
1|2	0.3902	1|12	0.0056	121	0.0813
2|2	0.6098	2|12	0.9944	122	0.0016
		1|21	0.2215	211	0.0144
		2|21	0.7785	212	0.0357
		1|22	0.0027	221	0.2859
		2|22	0.9973	222	0.5723
$\eta^{B\|A=1} = \eta^{B\|A=2} = 0.4878$		$\eta^{C\|A=1,B=1}$ =0.4831 $\eta^{C\|A=1,B=2}$ =0.0746 $\eta^{C\|A=2,B=1}$ =0.4152 $\eta^{C\|A=2,B=2}$ = 0.5723			

Naturalmente il modello causale viene stimato correttamente solo se si presuppone la gerarchia nel comando di stima e così facendo si ottiene

Tabella 19: stima sub-modello " primo caso con solo effetto diretto"

```
*** LOG-LINEAR PARAMETERS ***

* TABLE A [or P(A)] *

  effect          lambda  exp(lambda)
  A
   1             -1.3863   0.2500
   2              1.3863   4.0000
* TABLE AB [or P(B|A)] *

  effect          lambda  exp(lambda)
  B
   1             -0.2231   0.8000
   2              0.2231   1.2500
  AB
   1 1            0.0000   1.0000
   1 2            0.0000   1.0000
   2 1            0.0000   1.0000
   2 2            0.0000   1.0000

* TABLE CAB [or P(C|AB)] *

  effect          lambda  exp(lambda)
  C
   1             -1.6094   0.2000
   2              1.6094   5.0000
  CA
   1 1            0.1823   1.2000
   1 2           -0.1823   0.8333
   2 1           -0.1823   0.8333
   2 2            0.1823   1.2000
  CB
   1 1            1.1632   3.2000
   1 2           -1.1632   0.3125
   2 1           -1.1632   0.3125
   2 2            1.1632   3.2000
  CAB
   1 1 1          0.0000   1.0000
   1 1 2          0.0000   1.0000
   1 2 1          0.0000   1.0000
   1 2 2          0.0000   1.0000
   2 1 1          0.0000   1.0000
   2 1 2          0.0000   1.0000
   2 2 1          0.0000   1.0000
   2 2 2          0.0000   1.0000
```

```
*** (CONDITIONAL) PROBABILITIES ***

* P(A) *

  1               0.0588
  2               0.9412

* P(B|A) *

  1 | 1           0.3902
  2 | 1           0.6098
  1 | 2           0.3902
  2 | 2           0.6098

* P(C|AB) *

  1 | 1 1         0.3710
  1 | 1 2         0.0056
  1 | 2 1         0.2215
  1 | 2 2         0.0027
  2 | 1 1         0.6290
  2 | 1 2         0.9944
  2 | 2 1         0.7785
  2 | 2 2         0.9973
```

Nella tabella 19 (colonna 1) si osservano i parametri influenti nella condizionata P(C| A, B) nella “table CAB”, quelli influenti nella condizionata P(B|A) nella “table AB” e quelli influenti nella marginale di P(A) nella “table A”. Il parametro, la cui assenza nel processo generatore dei dati determina il sotto modello, è τ^{AB} che risulta identificato uguale ad 1. L’effetto di A su C (τ^{CA}) è 1.2, quello di B su C (τ^{CB}) è 3.2. L’interazione tra le variabili A,B e C, nulla nel processo generatore dei dati, è nulla ($\tau^{CAB} = 1 => \lambda^{CAB} = 0$).

1.4.4 B non influisce su C: secondo caso del solo effetto diretto

Ora si crea un modello in cui A influisce su C e B direttamente, ma B non influisce su C (equivalente ad una regressione multivariata). Si continuano a creare modelli gerarchici, perciò si costruiscono prima la probabilità di A e poi le 2 singole condizionate C|A e B|A. Questa volta le 2 condizionate sono indipendenti tra loro e sono allo stesso livello di gerarchia, quindi si può creare prima C|A e poi B|A o viceversa senza cambiare il risultato finale. Date le condizionate, si passa alla congiunta CAB semplicemente moltiplicando le 2 condizionate tra loro e la probabilità di A.

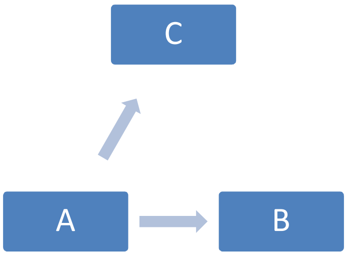


Figura 4: modello secondo caso del solo effetto diretto

I parametri moltiplicativi utilizzati sono:

$$\tau^{A} = 0.25$$

$$\tau^{B} = 0.8$$

$$\tau^{C} = 0.2$$
$$\tau^{AB} = 0.5$$

$$\tau^{AC} = 1.2$$

$$\tau^{CB} = 1$$

$$\tau^{CAB} = 1$$

Tabella 20: condizionata di B , di C e congiunta CAB, sub-modello "secondo caso con solo effetto diretto"

b|a		c|ab		cab	
1|1	0.1379	1|11	0.0545	111	0.0004
2|1	0.8621	2|11	0.9455	112	0.0028
1|2	0.7191	1|12	0.0545	121	0.0183
2|2	0.2809	2|12	0.9455	122	0.0072
		1|21	0.0270	211	0.0077
		2|21	0.9730	212	0.0479
		1|22	0.0270	221	0.6585
		2|22	0.9730	222	0.2572
$\eta^{B\|A=1}$ =0.3448 $\eta^{B\|A=2} = 0.4494$		$\eta^{C\|A=1,B=1} = \eta^{C\|A=1,B=2}$ =0.2269 $\eta^{C\|A=2,B=1} = \eta^{C\|A=2,B=2} = 0.1622$			

Stimando il modello in maniera causale, cioè A, B e C|AB, il risultato è:

Tabella 21: stima sub-modello "secondo caso con solo effetto diretto"

```
*** LOG-LINEAR PARAMETERS ***

* TABLE A [or P(A)] *

  effect           lambda  exp(lambda)
  A
   1              -1.3863   0.2500
   2               1.3863   4.0000

* TABLE AB [or P(B|A)] *

  effect           lambda  exp(lambda)
  B
   1              -0.2231   0.8000
   2               0.2231   1.2500
  AB
   1 1            -0.6931   0.5000
   1 2             0.6931   2.0000
   2 1             0.6931   2.0000
   2 2            -0.6931   0.5000

* TABLE CAB [or P(C|AB)] *

  effect           lambda  exp(lambda)
  C
   1              -1.6094   0.2000
   2               1.6094   5.0000
  CA
   1 1             0.1823   1.2000
   1 2            -0.1823   0.8333
   2 1            -0.1823   0.8333
   2 2             0.1823   1.2000
  CB
   1 1             0.0000   1.0000
   1 2             0.0000   1.0000
   2 1             0.0000   1.0000
   2 2             0.0000   1.0000
  CAB
   1 1 1           0.0000   1.0000
   1 1 2           0.0000   1.0000
   1 2 1           0.0000   1.0000
   1 2 2           0.0000   1.0000
   2 1 1           0.0000   1.0000
   2 1 2           0.0000   1.0000
   2 2 1           0.0000   1.0000
   2 2 2           0.0000   1.0000
```

```
*** (CONDITIONAL) PROBABILITIES ***

* P(A) *

  1              0.0588
  2              0.9412

* P(B|A) *

  1 | 1          0.1379
  2 | 1          0.8621
  1 | 2          0.7191
  2 | 2          0.2809

* P(C|AB) *

  1 | 1 1        0.0545
  1 | 1 2        0.0545
  1 | 2 1        0.0270
  1 | 2 2        0.0270
  2 | 1 1        0.9455
  2 | 1 2        0.9455
  2 | 2 1        0.9730
  2 | 2 2        0.9730
```

1.5 Effetti diretti , effetti indiretti e totali

Ora si studia l'effetto totale esistente tra A e C. Nelle sezioni precedenti si è utilizzato come metodo di identificazione l'"effect code", in questo paragrafo si utilizza inizialmente il "dummy code". L'analisi degli effetti può essere fatta anche con l'effect code, ma risulta semplificata con il dummy code a causa dell'uso del minor numero di parametri. Il dummy code ipotizza che una modalità sia la baseline e che le altre modalità vengano calcolate come differenza da questa. La baseline per tutte le variabili viene in questo caso considerata la modalità 1 e ciò porta a :

$\tau^{A=1}$=$\tau^{B=1} = \tau^{C=1}$ =1

Questo comporta:

$$\tau^{A=1B=1} = \tau^{A=1B=2} = \tau^{A=2B=1} = 1$$

$$\tau^{C=1B=1} = \tau^{C=1B=2} = \tau^{C=2B=1} = 1$$

$$\tau^{A=1C=1} = \tau^{A=1C=2} = \tau^{A=2C=1} = 1$$

Come affermato da Croon et al.(2009), l'effetto totale si può vedere dall'analisi della tabella AC e precisamente si calcola dalla probabilità condizionata di C dato A. Ricordando la scomposizione della condizionata (6), risulta

$\pi^{\,C=k|A=i} = \eta^{C|A=i}\mu^{C=k}\mu^{C=k\,A=i}$ con i,j =1,2

dove $\mu^{C=k}$ è l'effetto one-variable e $\mu^{C=k\,A=i}$ è l'effetto two-variables.
Per questo motivo $\mu^{C=k\,A=i}$ è l'effetto totale di A su C. Si è utilizzato il simbolo μ anziché il finora usato τ, per distinguere l'effetto totale da quello diretto. La presenza di altre variabili, oltre A e C, influenzano sia il parametro one-variable μ^{C} , attraverso quello che successivamente sarà chiamato effetto cella, sia il parametro della relazione di A su C μ^{CA}, attraverso l'effetto diretto, indiretto e cella. E' il solo studio del parametro μ^{CA} che è importante nell'analisi strutturale. Si noti che per i modelli causali la variabile risposta è la probabilità di trovarsi in una determinata cella e quindi i "regressori" influenzano proprio la probabilità.
Esistono 2 tipologie di effetti che influiscono sulla probabilità totale di trovarsi in una cella CA, β è l'effetto sulla probabilità condizionata di un aumento di A; α, invece, è l'effetto sulla probabilità condizionata di uno spostamento di C (il ruolo che α e β giocano sull'effetto totale e il loro significato è spiegato maggiormente nel dettaglio nell'appendice A.1). Quindi

$$\alpha = \frac{P(C=1|A=2)}{P(C=2|A=2)} = \frac{1}{\tau^{C=2}\tau^{A=2C=2}}\left[\frac{\eta^{C|A=2B=1} + \eta^{C|A=2B=2}\tau^{B=2}\tau^{B=2A=2}}{\eta^{C|A=2B=1} + \eta^{C|A=2B=2}\tau^{C=2\,B=2}\tau^{B=2}\tau^{B=2A=2}}\right]$$

$$\beta = \frac{P(C=2|A=1)}{P(C=2|A=2)}$$
$$= \frac{1}{\tau^{A=2C=2}}\left\{\frac{\eta^{B|A=1}}{\eta^{B|A=2}}\left[\frac{\eta^{C|A=1B=1} + \eta^{C|A=1B=2}\tau^{B=2}\tau^{B=2C=2}}{\eta^{C|A=2B=1} + \eta^{C|A=2B=2}\tau^{C=2\,B=2}\tau^{B=2}\tau^{B=2A=2}}\right]\right\}$$

dove

$$\eta^{C|A=1B=1} = \frac{1}{1+\tau^{C=2}}$$

$$\eta^{C|A=1B=2} = \frac{1}{1 + \tau^{C=2}\,\tau^{C=2\,B=2}}$$

$$\eta^{C|A=2\,B=1} = \frac{1}{1 + \tau^{C=2}\,\tau^{C=2\,A=2}}$$

$$\eta^{C|A=2\,B=2} = \frac{1}{1 + \tau^{C=2}\,\tau^{C=2\,A=2}\tau^{C=2\,B=2}}$$

L'effetto totale, definito $\mu^{C=2\,A=2}$, è una funzione di entrambi gli effetti:

$$\mu^{C=2\,A=2} = \frac{\alpha + 1 - \beta}{\alpha\beta}$$

Lo studio degli effetti diretti (τ^{AC}) e totali ($\mu^{C\,A}$) può essere effettuato anche nel caso dell'effect code. L' effetto totale continua ad essere funzione di α e β, ma questa volta si hanno 2 valori, uno il reciproco dell'altro, come vuole questa codificazione (dimostrazione nell'appendice A.2):

$$\mu^{C=2\,A=2} = \mu^{C=1A=1} = \sqrt[4]{\left(\frac{\alpha+1-\beta}{\alpha\beta}\right)} \quad \text{e} \quad \mu^{C=1\,A=2} = \mu^{C=2\,A=1} = \sqrt[4]{\left(\frac{\alpha\beta}{\alpha+1-\beta}\right)}$$

dove

$$\beta = (\tau^{A=1,C=2})^2 \left\{ (\tau^{A=2,B=1})^2 \frac{\eta^{B|A=1}}{\eta^{B|A=2}} \left[\frac{\eta^{C|A=1,B=1}(\tau^{B=1}\tau^{A=1\,B=1}\tau^{B=1\,C=2})^2 + \eta^{C|A=1,B=2}}{\eta^{C|A=2B=1}(\tau^{B=1}\tau^{A=2\,B=1}\tau^{B=1\,C=2})^2 + \eta^{C|A=2,B=2}} \right] \right\}$$

$$\alpha = (\tau^{A=2,C=1})^2(\tau^{C=1})^2 \left\{ (\tau^{C=2,B=1})^2 \frac{\eta^{C|A=2,B=1}(\tau^{B=1}\tau^{A=2\,B=1}\tau^{B=1\,C=1})^2 + \eta^{C|A=2,B=2}}{\eta^{C|A=2B=1}(\tau^{B=1}\tau^{A=2\,B=1}\tau^{B=1\,C=2})^2 + \eta^{C|A=2,B=2}} \right\}$$

La struttura di α e β, il cui significato è uguale nelle due codifiche, non cambia: β è composto dall'effetto diretto e dalla parte, che è compresa tra le parentesi graffe, dovuta alla presenza di B; α dall'effetto diretto, dall'effetto one-variable e dalla parte compresa tra parentesi graffe, che è quella dovuta alla presenza di B. L'effetto totale, grazie alla codificazione dummy, può anche essere così formulato (dimostrazione nell'appendice A.3):

$$\mu^{C=2\,A=2} = \tau^{A=2C=2} \left\{ \frac{\eta^{C|A=1B=1} + \eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=2B=1} + \eta^{C|A=2B=2}\tau^{B=2}\tau^{B=2A=2}} \left[\frac{\eta^{C|A=1B=1} + \eta^{C|A=1B=2}\tau^{B=2}\tau^{B=2C=2}}{\eta^{C|A=2B=1} + \eta^{C|A=2B=2}\tau^{C=2\,B=2}\tau^{B=2}\tau^{B=2A=2}} \right]^{-1} \right\}$$

dove l'effetto diretto è rappresentato da $\tau^{A=2C=2}$ e la parte fra parentesi è l'effetto indiretto e il cella, che sono definiti insieme effetto indiretto totale.

effetto_cella= $\frac{\eta^{C|A=1B=1}+\eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=1B=1}+\eta^{C|A=1B=2}\tau^{B=2}\tau^{B=2C=2}}\ \frac{\eta^{C|A=2B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{B=2C=2}}{\eta^{C|A=2B=1}+\eta^{C|A=2B=2}\tau^{B=2}}$

effetto_indiretto= effetto_ind_tot/effetto_cella

1.5.1 Modello completo

Si inizia l'analisi con un modello completo in cui A influisce direttamente sia su B che su C e B, a sua volta, influisce su C. Si è ancora in un modello causale e i dati vengono costruiti in maniera simile al modello causale completo con effect code. I parametri moltiplicativi utilizzati ora sono:

$$\tau^{A} = 1.5$$

$$\tau^{B} = 2$$
$$\tau^{AB} = 0.4$$

$$\tau^{C} = 0.6$$

$$\tau^{AC} = 0.8$$

$$\tau^{CB} = 0.5$$

$$\tau^{CAB} = 1$$

La scelta di non utilizzare i parametri dei modelli precedenti in parametrizzazione dummy è solo per semplicità computazionale.

Tabella 22: probabilità di A,condizionata di B , di C e congiunta CAB, modello completo

a		b|a		c|ab		cab	
1	0.4	1|1	0.3333	1|11	0.6250	111	0.0833
2	0.6	2|1	0.6667	2|11	0.3750	112	0.2051
		1|2	0.5556	1|12	0.7692	121	0.2252
		2|2	0.4444	2|12	0.2308	122	0.2151
				1|21	0.6757	211	0.05
				2|21	0.3243	212	0.0616
				1|22	0.8065	221	0.1081
				2|22	0.1935	222	0.0516
$\eta^{A} = 0.4$		$\eta^{B\|A=1}$ =0.3333 $\eta^{B\|A=2}$ =0.5556		$\eta^{C\|A=1B=1}$ =0.625 $\eta^{C\|A=1B=2}$ = 0.7692 $\eta^{C\|A=2B=1}$ = 0.6757 $\eta^{C\|A=2B=2}$ = 0.8065			

La stima del modello è

Tabella 23:stima modello completo

```
*** LOG-LINEAR PARAMETERS ***

* TABLE A [or P(A)] *

  effect          lambda  exp(lambda)
  A
   1              0.0000   1.0000
   2              0.4055   1.5000

* TABLE AB [or P(B|A)] *

  effect          lambda  exp(lambda)
  B
   1              0.0000   1.0000
   2              0.6931   2.0000
  AB
   1 1            0.0000   1.0000
   1 2            0.0000   1.0000
   2 1            0.0000   1.0000
   2 2           -0.9163   0.4000

* TABLE CAB [or P(C|AB)] *

  effect          lambda  exp(lambda)
  C
   1              0.0000   1.0000
   2             -0.5108   0.6000
  CA
   1 1            0.0000   1.0000
   1 2            0.0000   1.0000
   2 1            0.0000   1.0000
   2 2           -0.2231   0.8000
  CB
   1 1            0.0000   1.0000
   1 2            0.0000   1.0000
   2 1            0.0000   1.0000
   2 2           -0.6931   0.5000
  CAB
   1 1 1          0.0000   1.0000
   1 1 2          0.0000   1.0000
   1 2 1          0.0000   1.0000
   1 2 2          0.0000   1.0000
   2 1 1          0.0000   1.0000
   2 1 2          0.0000   1.0000
   2 2 1          0.0000   1.0000
   2 2 2          0.0000   1.0000
```

```
*** (CONDITIONAL) PROBABILITIES ***

* P(A) *

  1               0.4000
  2               0.6000

* P(B|A) *

  1 | 1           0.3333
  2 | 1           0.6667
  1 | 2           0.5556
  2 | 2           0.4444

* P(C|AB) *

  1 | 1 1         0.6250
  1 | 1 2         0.7692
  1 | 2 1         0.6757
  1 | 2 2         0.8065
  2 | 1 1         0.3750
  2 | 1 2         0.2308
  2 | 2 1         0.3243
  2 | 2 2         0.1935
```

Da questa tabella si possono calcolare i vari effetti:

β=1.0475

α=2.75655

Effetto_indiretto_tot=1.1727552=>log(effetto_indiretto_tot)=0.1593

Effetto_cella=1.0049=>log(effetto_cella)=0.0049

Effetto_indiretto=1.1670=>log(effetto_indiretto)=0.15442

Effetto_totale=0.938204 => log(effetto_totale)=-0.06379

L'effetto totale è maggiore (in termini assoluti minore) dell'effetto diretto -0.2231. Per l'effetto diretto, un aumento di A porta ad una diminuzione di C, per l'effetto indiretto un aumento di A porta ad una diminuzione di B, che comporta un aumento di C, quindi i 2 effetti sono tra loro opposti. Questo implica che uno neutralizzi parzialmente l'altro, infatti l'effetto indiretto totale ha segno opposto del diretto. L'effetto totale può anche essere stimato direttamente:

Tabella 24: stima effetto totale, modello completo

*** LOG-LINEAR PARAMETERS *** * TABLE A [or P(A)] * effect lambda exp(lambda) A 1 0.0000 1.0000 2 0.4055 1.5000 * TABLE CA [or P(C|A)] * effect lambda exp(lambda) C 1 0.0000 1.0000 2 -0.9502 0.3867 CA 1 1 0.0000 1.0000 1 2 0.0000 1.0000 2 1 0.0000 1.0000 2 2 -0.0638 0.9382	*** (CONDITIONAL) PROBABILITIES *** * P(A) * 1 0.4000 2 0.6000 * P(C|A) * 1 | 1 0.7212 1 | 2 0.7338 2 | 1 0.2788 2 | 2 0.2662

1.5.2 A non influisce su C: solo effetto indiretto

Si crea prima la probabilità di A, poi le condizionate B|A , successivamente la condizionata C|B e da queste la congiunta CAB. In questo caso la P(C|AB) può essere scritta come P(C|B), cioè la variabile C al netto di B è indipendente da A: A influisce su C solo indirettamente attraverso B. I parametri moltiplicativi usati per costruire i dati sono:

$$\tau^{A} = 1.5$$

$$\tau^{B} = 2$$

$$\tau^{C} = 0.6$$
$$\tau^{AB} = 0.4$$

$$\tau^{AC} = 1$$

$$\tau^{CB} = 0.5$$

$$\tau^{CAB} = 1$$

Tabella 25: probabilità di A,condizionata di B , di C e congiunta CAB, modello "solo effetto indiretto"

a		b|a		c|ab		cab	
1	0.4	1|1	0.3333	1|11	0.6250	111	0.0833
2	0.6	2|1	0.6667	2|11	0.3750	112	0.2051
		1|2	0.5556	1|12	0.7692	121	0.2083
		2|2	0.4444	2|12	0.2308	122	0.2051
				1|21	0.6250	211	0.05
				2|21	0.3750	212	0.0616
				1|22	0.7692	221	0.125
				2|22	0.2308	222	0.0616
$\eta^{A} = 0.4$		$\eta^{B\|A=1}$ =0.3333 $\eta^{B\|A=2}$ =0.5554		$\eta^{C\|A=1B=1}$ =0.625 $\eta^{C\|A=1B=2}$ =0.7692 $\eta^{C\|A=2B=1}$ =0.625 $\eta^{C\|A=2B=2}$ =0.7692			

La stima del modello è

Tabella 26:stima modello "solo effetto indiretto"

```
*** LOG-LINEAR PARAMETERS ***

* TABLE A [or P(A)] *

  effect          lambda   exp(lambda)
  A
   1              0.0000   1.0000
   2              0.4055   1.5000
* TABLE AB [or P(B|A)] *

  effect          lambda   exp(lambda)
  B
   1              0.0000   1.0000
   2              0.6931   2.0000
  AB
   1 1            0.0000   1.0000
   1 2            0.0000   1.0000
   2 1            0.0000   1.0000
   2 2           -0.9163   0.4000

* TABLE CAB [or P(C|AB)] *

  effect          lambda   exp(lambda)
  C
   1              0.0000   1.0000
   2             -0.5108   0.6000
  CA
   1 1            0.0000   1.0000
   1 2            0.0000   1.0000
   2 1            0.0000   1.0000
   2 2            0.0000   1.0000
  CB
   1 1            0.0000   1.0000
   1 2            0.0000   1.0000
   2 1            0.0000   1.0000
   2 2           -0.6931   0.5000
```

```
*** (CONDITIONAL) PROBABILITIES ***

* P(A) *

  1               0.4000
  2               0.6000

* P(B|A) *

  1 | 1           0.3333
  2 | 1           0.6667
  1 | 2           0.5556
  2 | 2           0.4444

* P(C|AB) *

  1 | 1 1         0.6250
  1 | 1 2         0.7692
  1 | 2 1         0.6250
  1 | 2 2         0.7692
  2 | 1 1         0.3750
  2 | 1 2         0.2308
  2 | 2 1         0.3750
  2 | 2 2         0.2308
```

```
  CAB
   1 1 1          0.0000   1.0000
   1 1 2          0.0000   1.0000
   1 2 1          0.0000   1.0000
   1 2 2          0.0000   1.0000
   2 1 1          0.0000   1.0000
   2 1 2          0.0000   1.0000
   2 2 1          0.0000   1.0000
   2 2 2          0.0000   1.0000
```

L’effetto totale questa volta è composto dall’effetto indiretto:

β=0.896907

α=2.216495

Effetto_ind_tot=1.1668

Effetto_cella=1

Effetto_indiretto=1.1668

Effetto_totale=1.1668 => log(effetto_totale)=0.154265

L’effetto totale è positivo, questo si spiega perché A↑ =>B↓=>C↑. I parametri di interesse dell’effetto indiretto sono $\tau^{A=2\ B=2}$ e $\tau^{B=2\ C=2}$(vedi Appendice A.3.3,A.4)

Tabella 27: stima effetto totale, modello “solo effetto indiretto”

```
*** LOG-LINEAR PARAMETERS ***

* TABLE A [or P(A)] *

  effect          lambda  exp(lambda)
  A
   1              0.0000   1.0000
   2              0.4055   1.5000
* TABLE CA [or P(C|A)] *

  effect          lambda  exp(lambda)
  C
   1              0.0000   1.0000
   2             -0.9502   0.3867
  CA
   1 1            0.0000   1.0000
   1 2            0.0000   1.0000
   2 1            0.0000   1.0000
   2 2            0.1543   1.1668
```

```
*** (CONDITIONAL) PROBABILITIES ***

* P(A) *

  1               0.4000
  2               0.6000

* P(C|A) *

  1 | 1           0.7212
  1 | 2           0.6891
  2 | 1           0.2788
  2 | 2           0.3109
```

1.5.3 A non influisce su B: solo effetto cella ed effetto diretto

Come per il modello con effect code, prima si creano le probabilità di A e di B e successivamente la condizionata C|AB. I parametri moltiplicativi usati per costruire i dati sono:

$$\tau^{A} = 1.5$$

$$\tau^{B} = 2$$

$$\tau^{C} = 0.6$$
$$\tau^{AB} = 1$$

$$\tau^{AC} = 0.8$$

$$\tau^{CB} = 0.5$$

$$\tau^{CAB} = 1$$

Tabella 28: probabilità di A,condizionata di B , di C e congiunta CAB, modello "effetto cella ed effetto diretto"

a		b|a		c|ab	
1	0.4	1|1	0.3333	1|11	0.6250
2	0.6	2|1	0.6667	2|11	0.3750
		1|2	0.3333	1|12	0.7692
		2|2	0.6667	2|12	0.2308
				1|21	0.6757
				2|21	0.3243
				1|22	0.8065
				2|22	0.1935
$\eta^{A} = 0.4$		$\eta^{B\|A=2} = \eta^{B\|A=1}$ =0.3333		$\eta^{C\|A=1B=1}$ =0.625 $\eta^{C\|A=1B=2} = 0.7692$ $\eta^{C\|A=2B=1} = 0.6757$ $\eta^{C\|A=2B=2} = 0.8065$	

cab	
111	0.0833
112	0.2051
121	0.1351
122	0.3226
211	0.05
212	0.0616
221	0.0649
222	0.0774

La stima del modello è

Tabella 29:stima modello “effetto cella ed effetto diretto”

```
*** LOG-LINEAR PARAMETERS ***

* TABLE A [or P(A)] *

  effect          lambda  exp(lambda)
  A
   1              0.0000   1.0000
   2              0.4055   1.5000
* TABLE AB [or P(B|A)] *

  effect          lambda  exp(lambda)
  B
   1              0.0000   1.0000
   2              0.6931   2.0000
  AB
   1 1            0.0000   1.0000
   1 2            0.0000   1.0000
   2 1            0.0000   1.0000
   2 2            0.0000   1.0000

* TABLE CAB [or P(C|AB)] *

  effect          lambda  exp(lambda)
  C
   1              0.0000   1.0000
   2             -0.5108   0.6000
  CA
   1 1            0.0000   1.0000
   1 2            0.0000   1.0000
   2 1            0.0000   1.0000
   2 2           -0.2231   0.8000
  CB
   1 1            0.0000   1.0000
   1 2            0.0000   1.0000
   2 1            0.0000   1.0000
   2 2           -0.6931   0.5000
  CAB
   1 1 1          0.0000   1.0000
   1 1 2          0.0000   1.0000
   1 2 1          0.0000   1.0000
   1 2 2          0.0000   1.0000
   2 1 1          0.0000   1.0000
   2 1 2          0.0000   1.0000
   2 2 1          0.0000   1.0000
   2 2 2          0.0000   1.0000
```

```
*** (CONDITIONAL) PROBABILITIES ***

* P(A) *

  1               0.4000
  2               0.6000

* P(B|A) *

  1 | 1           0.3333
  2 | 1           0.6667
  1 | 2           0.3333
  2 | 2           0.6667

* P(C|AB) *

  1 | 1 1         0.6250
  1 | 1 2         0.7692
  1 | 2 1         0.6757
  1 | 2 2         0.8065
  2 | 1 1         0.3750
  2 | 1 2         0.2308
  2 | 2 1         0.3243
  2 | 2 2         0.1935
```

Da questa tabella si possono calcolare i vari effetti:

β=1.17587

α=3.2169

Effetto_indiretto_totale=1.0049=> log(effetto_ind_totale)=0.0049

Effetto_cella=1.0049=>log(effetto_cella)=0.0049

Effetto_indiretto=1

Effetto_totale=0.8039 => log(effetto_totale)=-0.2182

In questo modello l'effetto totale è molto simile all'effetto diretto (-0.2231), perché A influisce solo direttamente su C, ma l'esservi B nella condizionata di C non permette ai 2 effetti (totale-parziale) di essere uguali. La causa è l'effetto cella dovuto all'interazione tra τ^{CB} e τ^{CA}, per la presenza delle costanti di normalizzazione (vedi Appendice A.3.2,A.4). L'effetto indiretto totale è composto dal solo effetto cella. Se per ipotesi non ci fosse l'effetto cella ($\tau^{CB} \longrightarrow 1,$ cioè quasi assenza di effetto di B su C) allora l'effetto totale sarebbe uguale all'effetto diretto (solo A influisce su C). Il parametro di interesse dell'effetto totale, comunque, rimane τ^{CA}; anche in questo caso se si stima la tabella AC risulta:

Tabella 30: stima effetto totale, modello "effetto cella ed effetto diretto"

<pre>*** LOG-LINEAR PARAMETERS ***
 	
* TABLE A [or P(A)] *
 	
  effect   lambda exp(lambda)
  A
   1   0.0000 1.0000
   2   0.4055 1.5000
 	
* TABLE CA [or P(C|A)] *
 	
  effect   lambda exp(lambda)
  C
   1   0.0000 1.0000
   2   -0.9502 0.3867
  CA
   1 1   0.0000 1.0000
   1 2   0.0000 1.0000
   2 1   0.0000 1.0000
   2 2   -0.2182 0.8039</pre>
	<pre>*** (CONDITIONAL) PROBABILITIES ***
	
	* P(A) *
	
	  1 0.4000
	  2 0.6000
	
	* P(C|A) *
	
	  1 | 1 0.7212
	  1 | 2 0.7629
	  2 | 1 0.2788
	  2 | 2 0.2371</pre>

1.5.4 B non influisce su C: solo effetto diretto

Si crea prima la probabilità di A, poi le condizionate B|A e C|A e da queste la congiunta CAB. I parametri moltiplicativi usati per costruire i dati sono:

$$\tau^{A} = 1.5$$

$$\tau^{B} = 2$$

$$\tau^{C} = 0.6$$
$$\tau^{AB} = 0.4$$

$$\tau^{AC} = 0.8$$

$$\tau^{CB} = 1$$

$$\tau^{CAB} = 1$$

Tabella 31: probabilità di A,condizionata di B , di C e congiunta CAB, modello "solo effetto diretto"

a	b|a		c|ab		cab	
1 0.4	1|1	0.3333	1|11	0.6250	111	0.0833
2 0.6	2|1	0.6667	2|11	0.3750	112	0.1667
	1|2	0.5556	1|12	0.6250	121	0.2252
	2|2	0.4444	2|12	0.3750	122	0.1802
			1|21	0.6757	211	0.05
			2|21	0.3243	212	0.1
			1|22	0.6757	221	0.1081
			2|22	0.3243	222	0.0865
$\eta^{A} = 0.4$	$\eta^{B\|A=1}$ =0.3333 $\eta^{B\|A=2}$ =0.5556		$\eta^{C\|A=1B=1}$ =0.625 $\eta^{C\|A=1B=2} = 0.625$ $\eta^{C\|A=2B=1} = 0.6757$ $\eta^{C\|A=2B=2} = 0.6757$			

La stima del modello è

Tabella 32:stima modello "solo effetto diretto"

```
*** LOG-LINEAR PARAMETERS ***

* TABLE A [or P(A)] *

  effect          lambda   exp(lambda)
  A
   1               0.0000   1.0000
   2               0.4055   1.5000

* TABLE AB [or P(B|A)] *

  effect          lambda   exp(lambda)
  B
   1               0.0000    1.0000
   2               0.6931    2.0000
  AB
   1 1             0.0000    1.0000
   1 2             0.0000    1.0000
   2 1             0.0000    1.0000
   2 2            -0.9163    0.4000

* TABLE CAB [or P(C|AB)] *

  effect          lambda   exp(lambda)
  C
   1               0.0000    1.0000
   2              -0.5108    0.6000
  CA
   1 1             0.0000    1.0000
   1 2             0.0000    1.0000
   2 1             0.0000    1.0000
   2 2            -0.2231    0.8000
  CB
   1 1             0.0000    1.0000
   1 2             0.0000    1.0000
   2 1             0.0000    1.0000
   2 2             0.0000    1.0000
```

```
*** (CONDITIONAL) PROBABILITIES ***

* P(A) *

 1               0.4000
  2               0.6000

* P(B|A) *

  1 | 1           0.3333
  2 | 1           0.6667
  1 | 2           0.5556
  2 | 2           0.4444

* P(C|AB) *

  1 | 1 1         0.6250
  1 | 1 2         0.6250
  1 | 2 1         0.6757
  1 | 2 2         0.6757
  2 | 1 1         0.3750
  2 | 1 2         0.3750
  2 | 2 1         0.3243
  2 | 2 2         0.3243
```

```
  CAB
   1 1 1          0.0000   1.0000
   1 1 2          0.0000   1.0000
   1 2 1          0.0000   1.0000
   1 2 2          0.0000   1.0000
   2 1 1          0.0000   1.0000
   2 1 2          0.0000   1.0000
   2 2 1          0.0000   1.0000
   2 2 2          0.0000   1.0000
```

Da questa tabella si possono calcolare i vari effetti:

β=1.15625

α=2.083333

Effetto_indiretto_tot=1 => log(effetto_ind_totale)=0

Effetto_cella=1

Effetto_indiretto=1

Effetto_totale=0.8 => log(effetto_totale)=-0.22314

In questo caso è solo A che influisce su B e, non essendoci alcuna influenza di B , sia diretta che indiretta sui parametri η, l'effetto totale è uguale all'effetto parziale.

$\beta = \frac{1+\tau^{C=2\ A=2}\tau^{C=2}}{\tau^{C=2\ A=2}(1+\tau^{C=2})}$

$\alpha = \frac{1}{\tau^{C=2\ A=2}\tau^{C=2}}$

Effetto totale=$\tau^{C=2\ A=2}$

Tabella 33: stima effetto totale, modello "solo effetto diretto"

```
*** LOG-LINEAR PARAMETERS ***

* TABLE A [or P(A)] *

  effect          lambda   exp(lambda)
  A
   1              0.0000   1.0000
   2              0.4055   1.5000

* TABLE CA [or P(C|A)] *

  effect          lambda   exp(lambda)
  C
   1              0.0000   1.0000
   2             -0.5108   0.6000
  CA
   1 1            0.0000   1.0000
   1 2            0.0000   1.0000
   2 1            0.0000   1.0000
   2 2           -0.2231   0.8000
```

Tutte le volte che α e β tendono ai valori visti sopra, l'effetto totale sarà uguale all'effetto diretto.

1.5.5 Effetto totale nullo

Ora si creano dei dati con effetto totale nullo: perché ciò avvenga bisogna che β sia uguale ad 1 (dimostrazione in Appendice A.5), cioè

$$\beta = 1 => = \frac{\eta^{B|A=1}}{\eta^{B|A=2}} \frac{1}{\tau^{A=2C=2}} \left[\frac{\eta^{C|A=1B=1} + \eta^{C|A=1B=2}\tau^{B=2}\tau^{B=2A=2}}{\eta^{C|A=2B=1} + \eta^{C|A=2B=2}\tau^{C=2\,B=2}\tau^{B=2}\tau^{B=2A=2}}\right] = 1$$

Questo è logico pensando a com'è costruito β:

$$\beta = \frac{P(C=2|A=1)}{P(C=2|A=2)} = 1 => P(C=2|A=1) = P(C=2|A=2)$$

cioè essere nella cella A=1 o A=2, non influisce sulla probabilità di trovarsi in C=2. Avendo variabili dicotomiche questa uguaglianza significa l'indipendenza tra A e C. Se, invece, si interpreta nella formulazione diretto- indiretto_totale, avere un effetto totale nullo vuol dire porre l'effetto indiretto uguale all'inverso dell'effetto diretto, cioè:

$$\tau^{A=2C=2} = \left\{\frac{\eta^{C|A=1B=1} + \eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=2B=1} + \eta^{C|A=2B=2}\tau^{B=2}\tau^{B=2A=2}} \left[\frac{\eta^{C|A=1B=1} + \eta^{C|A=1B=2}\tau^{B=2}\tau^{B=2C=2}}{\eta^{C|A=2B=1} + \eta^{C|A=2B=2}\tau^{C=2\,B=2}\tau^{B=2}\tau^{B=2A=2}}\right]^{-1}\right\}^{-1}$$

I dati sono creati come nel modello completo, cioè la probabilità di A, la condizionata B|A e successivamente la condizionata C|AB, ma si è ricercata una combinazione di $\tau^{AC}, \tau^{CB}\ e\ \tau^{AB}$ tale che β sia 1

$$\tau^{A} = 1.5$$

$$\tau^{B} = 2$$

$$\tau^{AB} = 0.71$$

$$\tau^{C} = 0.6$$

$$\tau^{AC} = 0.9$$

$$\tau^{CB} = 0.3$$

$$\tau^{CAB} = 1$$

Tabella 34: probabilità di A,condizionata di B , di C e congiunta CAB, modello "effetto totale nullo"

a		b|a		c|ab		cab	
1	0.4	1|1	0.3333	1|11	0.625	111	0.0833
2	0.6	2|1	0.6667	2|11	0.375	112	0.2260
		1|2	0.4132	1|12	0.8475	121	0.1610
		2|2	0.5868	2|12	0.1525	122	0.3030
				1|21	0.6494	211	0.05
				2|21	0.3506	212	0.0406
				1|22	0.8606	221	0.0870
				2|22	0.1394	222	0.0491
$\eta^{A} = 0.4$		$\eta^{B\|A=1}$ =0.3333 $\eta^{B\|A=2}$ =0.4132		$\eta^{C\|A=1B=1}$ =0.625 $\eta^{C\|A=1B=2} = 0.8475$ $\eta^{C\|A=2B=1} = 0.6494$ $\eta^{C\|A=2B=2} = 0.8606$			

La stima del modello è

Tabella 35:stima modello "effetto totale nullo"

```
*** LOG-LINEAR PARAMETERS ***

* TABLE A [or P(A)] *

  effect          lambda   exp(lambda)
  A
   1              0.0000   1.0000
   2              0.4055   1.5000

* TABLE AB [or P(B|A)] *

  effect          lambda   exp(lambda)
  B
   1              0.0000    1.0000
   2              0.6931    2.0000
  AB
   1 1            0.0000    1.0000
   1 2            0.0000    1.0000
   2 1            0.0000    1.0000
   2 2           -0.3425    0.7100

* TABLE CAB [or P(C|AB)] *

  effect          lambda   exp(lambda)
  C
   1              0.0000    1.0000
   2             -0.5108    0.6000
  CA
   1 1            0.0000    1.0000
   1 2            0.0000    1.0000
   2 1            0.0000    1.0000
   2 2           -0.1054    0.9000
  CB
   1 1            0.0000    1.0000
   1 2            0.0000    1.0000
   2 1            0.0000    1.0000
   2 2           -1.2040    0.3000
```

```
*** (CONDITIONAL) PROBABILITIES ***

* P(A) *

  1               0.4000
  2               0.6000

* P(B|A) *

  1 | 1           0.3333
  2 | 1           0.6667
  1 | 2           0.4132
  2 | 2           0.5868

* P(C|AB) *

  1 | 1 1         0.6250
  1 | 1 2         0.8475
  1 | 2 1         0.6494
  1 | 2 2         0.8606
  2 | 1 1         0.3750
  2 | 1 2         0.1525
  2 | 2 1         0.3506
  2 | 2 2         0.1394
```

```
CAB
   1 1 1          0.0000   1.0000
   1 1 2          0.0000   1.0000
   1 2 1          0.0000   1.0000
   1 2 2          0.0000   1.0000
   2 1 1          0.0000   1.0000
   2 1 2          0.0000   1.0000
   2 2 1          0.0000   1.0000
   2 2 2          0.0000   1.0000
```

L'effetto totale è nullo:

β=0.999969~1

α=3.411

Effetto_indiretto_tot=1.1111 => log(effetto_ind_totale)=0.1504

Effetto_cella=1.0064=>log(effetto_cella)=0.0064

Effetto_indiretto=1.1040=>log(effetto_ind)=0.0989

Effetto_totale=1 => log(effetto_totale)=0

L'effetto totale nullo può essere così spiegato: un aumento di A porta ad una diminuzione di B. La diminuzione di B fa aumentare C, cioè l'effetto indiretto di A su C è positivo. Questo è contrastato dall'effetto diretto di A su C che è negativo. I 2 effetti, essendo uno l'opposto dell'altro, portano ad avere un effetto totale nullo.

Tabella 36: stima effetto totale, modello "effetto totale nullo"

```
*** LOG-LINEAR PARAMETERS ***

* TABLE A [or P(A)] *

  effect          lambda   exp(lambda)
  A
   1              0.0000   1.0000
   2              0.4055   1.5000

* TABLE CA [or P(C|A)] *

  effect          lambda   exp(lambda)
  C
   1              0.0000   1.0000
   2             -1.2271   0.2932
  CA
   1 1            0.0000   1.0000
   1 2            0.0000   1.0000
   2 1            0.0000   1.0000
   2 2            0.0000   1.0000
```

```
*** (CONDITIONAL) PROBABILITIES ***

* P(A) *

  1               0.4000
  2               0.6000

* P(C|A) *

  1 | 1           0.7733
  1 | 2           0.7733
  2 | 1           0.2267
  2 | 2           0.2267
```

1.5.6 Effetto cella su $\tau^{C=2\ A=2}$

Come indicato nel modello "A non influisce su B", l'assenza di un legame tra A su B porta a un effetto totale molto simile a $\tau^{C=2\ A=2}$, ma non del tutto uguale a causa dell'effetto cella. Ma cos'è quest'ultimo? Esso deriva dalla costruzione dei dati , cioè dall'influenza diretta sia di A che di B su C e precisamente dai 4 η che sono presenti nella distribuzione di C|AB; affinché l'effetto totale sia uguale a quello diretto dovremmo avere

$\eta^{C|A=1B=1} = \eta^{C|A=1B=2}$ e $\eta^{C|A=2B=1} = \eta^{C|A=2B=2}$

Uguaglianze che avvengono solo se $\tau^{C=2\ B=2}$=1 (B non influisce su C), perciò al tendere di tale parametro ad 1, l'effetto cella diminuisce. Un altro caso in cui l'effetto cella su $\tau^{C=2\ A=2}$ è nullo, mentre rimane su $\tau^{C=2}$, è quando:

$\eta^{C|A=1B=1} = \eta^{C|A=2B=1}$ e $\eta^{C|A=1B=2} = \eta^{C|A=2B=2}$

dimostrando che l'annullamento dell'effetto cella avviene solo se si riescono ad avere 2 η differenti, cioè che solo 1 variabile influisce direttamente su C. Un'attenuazione dell'effetto cella può avvenire con la diminuzione degli η da 4 a 3? Si prova diminuendo le η. Questa diminuzione non è dovuta all'assenza di un effetto diretto, ma ad uguaglianze tra gli η ricavate da uguaglianze tra gli effetti diretti $\tau^{C=2\ A=2}$ e $\tau^{C=2\ B=2}$

1° modo: $\eta^{C|A=1B=1} = \eta^{C|A=2B=2} => \tau^{C=2\ A=2} = (\tau^{C=2\ B=2})^{-1}$

2° modo: $\eta^{C|A=1B=2} = \eta^{C|A=2B=1} => \tau^{C=2\ A=2} = \tau^{C=2\ B=2}$

e non c'è alcun miglioramento, l'effetto cella rimane perché A e B continuano ad influire direttamente su C. L'effetto cella, quindi, non dipende propriamente da η, ma dalla presenza/assenza degli effetti diretti che influiscono anche su η.

Tabella 35: ricerca della diminuzione dell'effetto cella: modello

modello di partenza	1° modo	2° modo
$\tau^{A} = 1.5$ $\tau^{B} = 2$ $\tau^{C} = 0.6$ $\tau^{AB} = 1$ $\tau^{CA} = 0.5$ $\tau^{CB} = 0.8$	$\tau^{C=2\ B=2} = (\tau^{C=2\ A=2})^{-1}$	$\tau^{C=2\ B=2} = \tau^{C=2\ A=2}$
$\eta^{C\|A=1B=1} = 0.6250$ $\eta^{C\|A=1B=2} = 0.6757$ $\eta^{C\|A=2B=1} = 0.7692$ $\eta^{C\|A=2B=2} = 0.8055$ β=1.1656 α=3.8554 Effetto_totale=0.5007	$\eta^{C\|A=1B=1} =0.6250$ $\eta^{C\|A=1B=2} =0.4545$ $\eta^{C\|A=2B=1} = 0.7692$ $\eta^{C\|A=2B=2} = 0.625$ β=1.494 α=2.0588 Effetto_totale=0.5038	$\eta^{C\|A=1B=1} =0.6250$ $\eta^{C\|A=1B=2} = 0.7692$ $\eta^{C\|A=2B=1} = 0.7692$ $\eta^{C\|A=2B=2} = 0.8696$ β=1.7015 α=5.1020 Effetto_totale=0.5068

L'effetto cella è quindi l'effetto "interazione" tra i 2 parametri effetti diretti $\tau^{C=2\ B=2} e\ \tau^{C=2\ A=2}$. Com'è ampiamente approfondito nell'Appendice A.3.2, la trasformazione logaritmica dell'effetto cella è sempre di segno opposto rispetto a $\ln(\tau^{C=2\ A=2})$ ed all'aumentare di $\tau^{C=2\ A=2}$ l'effetto cella diminuisce.

1.5.7 Conclusioni

In questa sezione è stato analizzato un modello causale a tre variabili A, B e C e i suoi sub-modelli. La parte innovativa di questo lavoro consiste nell'aver trovato un metodo per riuscire a determinare non solo l'effetto totale di A su C, ma anche nello scomporre l'effetto totale in effetto diretto e indiretto, suddivisione che è importante nei modelli causali in generale. Ci si potrebbe chiedere come mai l'analisi principale sia stata svolta sul parametro $\mu^{C=2\ A=2}$ e non sul parametro $\mu^{C=2}$ che pure influisce nella tabella di frequenza AC. Questo è spiegato ricordando che solo $\mu^{C=2\ A=2}$ misura l'effetto di A su C, mentre il parametro $\mu^{C=2}$ misura la variazione della probabilità $\pi^{C=2\ A=2}$al variare di C. Il parametro $\mu^{C=2}$ dipende dal solo effetto cella e quindi si differenzia da $\tau^{C=2}$ solo per come sono costruiti i dati.

Capitolo 2:
Structural equation models

2.1 Introduzione

I modelli SEM, cioè structural equation models, si usano per studiare congiuntamente sia se le variabili osservate siano indicatori di un numero di variabili latenti, attraverso l'analisi fattoriale, sia la direzione dei legami fra queste, attraverso l'analisi strutturale. Nella più classica metodologia SEM tutte le variabili, essendo ipotizzate distribuite normalmente, sono a media nulla e quindi le variabili del dataset iniziale devono essere trasformate in variabili "scarto dalla media".

2.1.1 Analisi fattoriale esplorativa

Nell'analisi fattoriale si presume che il legame tra le variabili che compongono un dataset sia dovuto a variabili non direttamente osservate. Per esempio se si dispone di un dataset composto da 2 variabili A e B tra loro correlate, l'analisi fattoriale ipotizza che questo legame sia dovuto ad una terza variabile C non presente nel datatset (e quindi latente) che influisce sia su A che su B. L'analisi fattoriale è o confermativa (CFA) o esplorativa (EFA). In entrambe le analisi le variabili osservate X, seguendo la classica simbologia SEM, sviluppata da Jöreskog (1973, 1977), Wiley (1973) e Keesling (1972), sono funzione delle variabili latenti ξ e di un termine d'errore δ. Il modello che lega tra loro le variabili latenti ξ e le variabili osservate X è lineare :

$$\begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \cdots & \lambda_{1l} \\ \vdots & \ddots & \vdots \\ \lambda_{n1} & \cdots & \lambda_{nl} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_l \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_n \end{bmatrix} = \Lambda_x \xi + \underline{\delta} \qquad (1)$$

dove

$\xi \sim N(0, \Phi)$

$\delta \sim N(0, \Theta_\delta)$

I parametri λ , chiamati loadings, rappresentano quanto le variabili latenti influiscano sulle variabili osservate, Φ è la matrice di varianza –covarianza delle variabili latenti ξ e Θ_δ è la matrice di varianza-covarianza dei termini d'errore ε.

Nell'analisi fattoriale esplorativa si specifica solo il numero di variabili latenti, denominate fattori, e quali variabili osservate siano da prendere in considerazione, ma non si fanno ipotesi sulle relazioni tra le variabili del modello. Nella formulazione più semplice di tale analisi, tutti i fattori sono tra loro incorrelati, incorrelati con gli errori e influiscono direttamente sulle variabili osservate, chiamate indicatori. La (1) diviene:

$$\begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \dots & \lambda_{1l} \\ \vdots & \ddots & \vdots \\ \lambda_{n1} & \dots & \lambda_{nl} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_l \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_n \end{bmatrix} \tag{2}$$

dove

$$\Phi = \begin{bmatrix} \phi_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \phi_{ll} \end{bmatrix}$$

$$\Theta_\delta = \begin{bmatrix} \theta_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \theta_{nn} \end{bmatrix}$$

L'analisi fattoriale esplorativa (Kim e Muller 1994) si sviluppa in tre stadi: nel primo passaggio si calcola la matrice di varianza – covarianza campionaria; nel secondo si determina il numero di fattori ipotizzati incorrelati; nel terzo si suddividono gli indicatori tra i fattori e se quest'ultimi sono correlati o non correlati. Esistono vari metodi per estrarre il numero dei fattori, i due principali sono quello di massima verosimiglianza(ML) e quello dei minimi quadrati, la cui variante è il MINRES. In entrambi i metodi, per vedere quale modello si adatti meglio ai dati, è possibile un test ${\chi^2}_{d_f}$ come consigliato da Bartholomew (2008):

$$U = N\{ln|FF' + U^2| - ln|R| + tr(R(FF' + U^2)^{-1}) - n\}$$

dove F è il vettore dei loadings, U^2 è la matrice diagonale delle "unique variance", N la grandezza del campione e n il numero delle variabili. I gradi di libertà si calcolano con la seguente formula:

$$d_f = \frac{(n-k)^2 - (n+k)}{2}$$

dove k è il numero di fattori ipotizzati. Questo test indica se i dati deviino significativamente dal modello a k fattori. Si continua poi a diminuire il numero di fattori finché il modello non devii significativamente dai dati. Trovato il numero di fattori, si attribuiscono ai fattori i propri indicatori. Se i fattori sono ipotizzati incorrelati si usa il metodo Varimax, se sono correlati il direct Oblim o il Promax. I due metodi calcolano i loadings dei fattori sulle variabili osservate: i loadings più alti significano che quel fattore influisce su quella variabile osservata, i loadings più bassi significano che quel fattore non influisce.

2.1.2 Analisi fattoriale confermativa

Nell'analisi fattoriale confermativa, oltre a specificare il numero dei fattori, si pongono vincoli motivati, i quali possono riguardare sia le correlazioni sia il legame fattori-indicatori (per esempio un fattore è legato a solo 2 indicatori). Ad esempio con indicatori congenerici, si può ipotizzare un modello in cui ogni variabile osservata è indicatore di un solo fattore. I fattori sono tutti correlati fra loro. In una CFA come quella del modello supposto, la relazione (1) tra variabili osservate e latenti diviene:

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} \lambda_{11} & 0 & \dots & 0 \\ \lambda_{21} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_{nl} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_l \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_n \end{bmatrix} \qquad (3)$$

dove

$$\Phi = \begin{bmatrix} \phi_{11} & \phi_{21} & \cdots & \phi_{1l} \\ \phi_{21} & \phi_{22} & \dots & \phi_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{l1} & \phi_{l2} & \cdots & \phi_{ll} \end{bmatrix}$$

$$\Theta_\delta = \begin{bmatrix} \theta_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \theta_{nn} \end{bmatrix}$$

Dopo aver determinato i modelli con le 2 analisi fattoriali (CFA e EFA), si cerca capire se questi sono identificati. L'identificazione richiede che i parametri del modello siano determinati univocamente a partire dalla matrice di covarianza empirica dal dataset.
Nell'EFA questo problema è risolto, ad esempio come affermato da Long (1994) rifacendosi a Lawley e Maxwell (1971), imponendo che il prodotto $\Lambda_x \Theta_\delta^{-1} \Lambda_x'$ abbia come risultato una matrice diagonale i cui elementi siano positivi, distinti e ordinati.
Nella CFA il problema è risolto imponendo vincoli basati su considerazioni sostanziali. Bollen (1989) fornisce una condizione necessaria e due sufficienti per studiare l'identificazione nell'analisi fattoriale confermativa. La condizione necessaria ma non sufficiente affinché il modello sia identificato è la t-rule: il numero di parametri da stimare, denominato t, deve essere minore od uguale al numero di elementi indipendenti della matrice di varianza-covarianza delle X, cioè t≤(1/2)(n)(n+1). Per esempio, se le X

sono 5, la matrice di varianza-covarianza avrà 15 elementi indipendenti, cioè le 5 varianze e le 10 covarianze. I parametri da stimare dovranno essere al massimo 15.
Una condizione sufficiente è la regola dei 3 indicatori. Essa richiede che ci sia un elemento diverso da 0 per ogni riga di Λ_x, 3 o più indicatori per fattori e Θ_δ sia diagonale. Un'altra condizione sufficiente è la regola dei 2 indicatori. Essa è applicabile in due casi: quando tutte le variabili latenti sono correlate e quando almeno una coppia è correlata. In entrambi i casi è sufficiente avere 2 o più indicatori, Θ_δ diagonale e un elemento diverso da zero per ogni riga di Λ_x affinché il modello sia identificato. Jöreskog e Sörbom hanno affermato che se la matrice di informazione è definita positiva si è certi che il modello sia identificato.
Un modello fattoriale non può essere identificato finché non viene data una "scala" alle variabili latenti. Questa "misura" è solitamente data in due modi: fissando la varianza dei fattori o un parametro λ per ogni fattore. In generale si sceglie di porre uguale all'unità o la varianza o il λ, in quest'ultimo modo la variabile latente ha la stessa scala della variabile osservata a cui si riferisce λ. Ad esempio, si suppone di avere 2 fattori e 4 variabili osservate. Per fissare la scala si pone o Var(ξ_1)= Var(ξ_2)=1 o $\lambda_{11} = \lambda_{22} = 1$, dove λ_{11} è il parametro di "loading" della variabile latente ξ_1 sulla variabile osservata X_1 e λ_{22} è il parametro di "loading" della variabile latente ξ_2 sulla variabile osservata X_2 .

2.1.3 Analisi strutturale

Nell'analisi fattoriale le relazioni tra i singoli fattori sono spiegate attraverso la correlazione, ma ciò non è sufficiente per spiegare un modello. Per questo motivo si ricorre alla parte strutturale, dove si dà una direzione alle relazioni tra fattori. Essa ha la tipica forma di una regressione, cioè è formata da una variabile risposta, che rappresenta l'evento da studiare, e da più regressori, che sono le variabili "causanti" l'evento studiato. Un modello strutturale è di due tipi: a variabili latenti o a variabili osservate. In un modello con tutte le variabili osservate, sono definite X le esogene, cioè quelle che non sono spiegate dal modello, e Y le endogene, cioè quelle spiegate dal modello (Jöreskog 1973, 1977, Wiley 1973 e Keesling 1972). La relazione è quindi del tipo:

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_s \end{bmatrix} = \begin{bmatrix} 0 & \beta_{12} & \cdots & \beta_{1s} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{s1} & \cdots & \cdots & 0 \end{bmatrix} \begin{bmatrix} Y_1 \\ \vdots \\ Y_s \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \cdots & \gamma_{1n} \\ \vdots & \ddots & \vdots \\ \gamma_{s1} & \cdots & \gamma_{sn} \end{bmatrix} \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_s \end{bmatrix} = \mathrm{B}\underline{Y} + \Gamma\underline{X} + \underline{\zeta}$$

(4)

dove

$\zeta \sim N(0, \Psi)$

La Ψ rappresenta la matrice di varianza-covarianza dei termini d'errore ζ.
In un modello con variabili latenti non si studia direttamente la relazione tra le variabili osservate X e Y, ma quella tra le latenti η e ξ. Le η e ξ rappresentano rispettivamente le

latenti endogene, spiegate dal modello strutturale, e quelle esogene, non spiegate dal modello strutturale. Le variabili osservate X sono gli indicatori delle variabili ξ strutturalmente esogene, le variabili Y sono indicatori delle variabili η strutturalmente endogene. Con l'introduzione delle latenti la relazione (4) diviene:

$$\begin{bmatrix} \eta_1 \\ \vdots \\ \eta_m \end{bmatrix} = \begin{bmatrix} 0 & \beta_{12} & \dots & \beta_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{m1} & \dots & \dots & 0 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_m \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \dots & \gamma_{1l} \\ \vdots & \ddots & \vdots \\ \gamma_{m1} & \dots & \gamma_{ml} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_l \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_m \end{bmatrix} = \mathrm{B}\underline{\eta} + \Gamma\underline{\xi} + \underline{\zeta} \qquad (5)$$

dove

$\xi \sim N(0, \Phi)$

$\zeta \sim N(0, \Psi)$

Ψ è la matrice di varianza-covarianza dell'errore ζ e Φ è la matrice di varianza-covarianza delle variabili latenti esogene ξ.
L'interpretazione della relazione tra le variabili osservate X e Y (4) e della relazione tra le variabili latenti ξ e η (5) è uguale. I parametri β rappresentano la relazione tra le variabili endogene, i parametri γ quelle tra le variabili esogene e le endogene. Per esempio nella (5), β_{1m} rappresenta come varia η_1 al variare di η_m e γ_{1l} rappresenta la variazione di η_1 al variare di ξ_l. Nella (5), Il termine d'errore è la parte non spiegata delle η, nella (4) la parte non spiegata delle Y.
Come nelle regressioni canoniche, il termine d'errore non è correlato con i regressori. Dalle relazioni appena viste, si capisce l'altra denominazione data al SEM, cioè quella di "modelli a equazioni simultanee": infatti si è in presenza di più equazioni che devono essere stimate simultaneamente perché alcuni regressori di un'equazione sono le variabili risposta delle altre.
Dopo aver determinato i modelli strutturali a variabili osservate e a variabili latenti, si verifica se questi sono identificati. Bollen (1989) fornisce "condizioni necessarie e/o sufficienti" per studiare l'identificazione in entrambi i modelli.
Si definisce l'equivalente della relazione (2) tra le osservate X e le latenti ξ, per le osservate Y e le latenti η, che è

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_s \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \dots & \lambda_{1m} \\ \vdots & \ddots & \vdots \\ \lambda_{s1} & \dots & \lambda_{sm} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_s \end{bmatrix} = \underline{Y} = \Lambda_y \underline{\eta} + \underline{\varepsilon} \qquad (2.a)$$

dove

$\varepsilon \sim N(0, \Theta_\varepsilon)$

Θ_ε è la matrice delle varianze- covarianze degli errori ε.

Si inizia con le regole di identificazione per un modello strutturale a variabili osservate, cioè quando Y≡η e X≡ξ. La condizione necessaria ma non sufficiente affinché il modello sia identificato è la t-rule: il numero di parametri da stimare, denominato t, deve essere minore od uguale al numero di elementi indipendenti della matrice di varianza-covarianza delle XY, cioè t≤(1/2)(m+l)(m+l+1). Una condizione sufficiente può essere la regola di B nulla, che richiede che la matrice dei β (cioè B) sia la matrice nulla, quindi di trovarsi in un semplice contesto di regressione multivariata. Un'altra condizione sufficiente è la regola ricorsiva, che richiede che B sia triangolare e che gli errori ζ siano incorrelati, cioè che la matrice delle varianze-covarianze degli errori, denominata Ψ, sia diagonale.
Ora si passa alle regole di identificazione per un modello strutturale a variabili latenti (Y≠η e X≠ξ). La t-rule continua a valere, cioè t≤(1/2)(m+l)(m+l+1). La condizione sufficiente è la regola dei 2 steps. Il primo step consiste nel riformulare il modello in termini di misura e di studiare l'identificazione in esso con le regole dell'analisi fattoriale confermativa. Il secondo step consiste nell'ipotizzare che le variabili strutturali siano osservate direttamente e nel controllare l'identificazione in un modello così. Per esempio si ipotizzi un modello in cui si hanno 2 X e 2 Y che indicano rispettivamente una ξ e una η, cioè si è in un modello così:

$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} \lambda_{1X} \\ \lambda_{2X} \end{bmatrix} \xi + \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix}$ e $\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} \lambda_{1Y} \\ \lambda_{2Y} \end{bmatrix} \eta + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$ con gli errori tra di loro incorrelati

$$\eta = \gamma_{11}\xi + \zeta$$

Con il primo step si controlla l'identificazione in un modello CFA del tipo:

$\begin{bmatrix} X_1 \\ X_2 \\ Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{21} & 0 \\ 0 & \lambda_{32} \\ 0 & \lambda_{42} \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix}$ con ξ e η correlati e gli errori δ incorrelati

Questo modello per la regola dei 2 indicatori è identificato, infatti le 2 latenti sono correlate, Θ_δ è diagonale e c'è un elemento diverso da 0 per ogni riga di Λ_x.
Con il secondo step si controlla l'identificazione nel modello strutturale ipotizzando che η e ξ siano direttamente osservate. Il modello è identificato per la regola della matrice B nulla. Con la regola dei 2 step si è visto che il modello è identificato.

2.2 SEM: stima

Un'ulteriore denominazione del modello SEM è "analisi della struttura della covarianza" ed è infatti attraverso la matrice di covarianza campionaria (chiamata **S**) che si ottengono i parametri e, precisamente, cercando di uguagliarla a quella definita nel modello (chiamata **Σ**):

$\mathbf{S}= \Sigma_{ipotizzata\ dal\ modello}$

dove

$\mathbf{S}=\begin{bmatrix} COV(\underline{Y}\ \underline{Y}') & COV(\underline{Y}\ \underline{X}') \\ COV(\underline{X}\ \underline{Y}') & COV(\underline{X}\ \underline{X}') \end{bmatrix}$

In generale è impossibile uguagliare le 2 matrici e quindi nel processo di stima si cercherà di minimizzare la distanza tra le 2. Rifacendosi a Bollen (1989), la funzione che misura la distanza tra **S** e **Σ** è la "fitting function", che viene così denotata F(**S Σ**). Una funzione per essere una fitting fuction deve possedere le seguenti caratteristiche:

1) F(**S Σ**) è scalare
2) F(**S Σ**) ≥ 0
3) F(**S Σ**)=0 se e solo se **S**= **Σ**
4) F(**S Σ**) è continua in **S** e **Σ**

Esistono 3 comuni fitting function e sono:

1)la "Unweighted least squares"

$F_{ULS}(\mathbf{S}\ \Sigma)=\mathrm{tr}[(\mathbf{S}-\ \Sigma)^2]$

2)la "Generalized least squares"

$F_{GLS}(\mathbf{S}\ \Sigma)=\mathrm{tr}[(\mathbf{S}-\ \Sigma)\mathbf{S}^{-1}]^2$

3)la "Maximum Likelihood"

$F_{ML}(\mathbf{S}\ \Sigma)=\mathrm{tr}\big(\mathbf{S}\ \Sigma^{-\mathbf{1}}\big)+[log|\Sigma|-log|\mathbf{S}|]-n$

Il metodo ULS soffre della dipendenza dalla scala, cioè le stime variano se si cambia l'unità di misura. I metodi GLS e ML, invece, sono indipendenti dalla scala usata.

Per ogni tipo di modello esiste la sua specifica **Σ** e quindi si deve definirne una per l'analisi fattoriale, una per il modello strutturale a variabili osservate e una per quello a variabili latenti. Per la determinazione analitica di **Σ** ci si rifà a Bollen (1989)

- **Determinazione Σ per l'analisi fattoriale confermativa**

Nell'analisi fattoriale confermativa le variabili osservate sono X e le latenti sono ξ, perciò S sarà la matrice delle varianze-covarianze osservate delle X. Nei modelli SEM non si utilizzano direttamente le variabili osservate, ma i loro scarti dalla media (cioè X=$X_{oss}-\bar{X}$, con $\bar{X}$ media campionaria) e quindi COV(X X') è uguale a E(X X'). Ricordando la relazione (2) tra fattori ed indicatori:

$\underline{X}=\Lambda_x\underline{\xi}+\delta$

con semplici passaggi si ottiene la matrice di varianza-covarianza del modello che diviene:

Σ= E($\underline{X}$ $\underline{X'}$)= $\Lambda_x \Phi \Lambda_x' + \Theta_\delta$.

- **Determinazione Σ per l'analisi strutturale a variabili osservate**

Nel modello strutturale a variabili osservate, cioè Y≡η e X≡ξ, le variabili osservate sono divise tra X e Y, perciò **S** è la matrice di varianza-covarianze delle XY. Le variabili osservate non vengono utilizzate direttamente, ma come scarti dalla media, cioè X=$X_{oss} - \bar{X}$ e Y=$Y_{oss} - \bar{Y}$. Per calcolare **Σ**, riformulando la relazione (4) tra le variabili osservate endogene e quelle osservate esogene, si ha:

$\underline{Y}$ =B $\underline{Y}$+Γ$\underline{X}$+$\underline{\zeta}$ => (I-B) $\underline{Y}$= Γ $\underline{X}$+$\underline{\zeta}$ => $\underline{Y}$= $(I - B)^{-1}$(Γ $\underline{X}$+$\underline{\zeta}$)

Con semplici passaggi si ottiene la matrice di varianza-covarianza delle Y che diviene:

E(X'Y)= $[(I - B)^{-1}(\Gamma \mathrm{E(X\,X')} \Gamma' + \Psi)(I - B)^{-1\prime}]$

La matrice campionaria **S** è formata dalle matrici di varianza-covarianza delle X, delle Y e dalle covarianze tra le X e le Y e quindi si calcola l'equivalente covarianza del modello

E(X'Y)= $\Phi\Gamma'[(\mathrm{I} - \mathrm{B})^{-1}]'$

La matrice risultante **Σ** è:

$$\mathbf{\Sigma}=\begin{bmatrix} [(I - B)^{-1}(\Gamma \mathrm{E(X\,X')} \Gamma' + \Psi)(I - B)^{-1\prime}] & [(\mathrm{I} - \mathrm{B})^{-1}]\Gamma \mathrm{E(X\,X')} \\ \mathrm{E(X\,X')}\Gamma'[(\mathrm{I} - \mathrm{B})^{-1}]' & \mathrm{E(X\,X')} \end{bmatrix}$$

- **Determinazione Σ per l'analisi strutturale a variabili latenti**

Nel modello strutturale a variabili latenti, per determinare **Σ** ,si unisce l'analisi fattoriale a quella del modello strutturale a variabili osservate.

Per definire E($\underline{Y}$ $\underline{Y'}$), si calcola COV($\underline{\eta}\, \underline{\eta}'$) che è uguale a E($\underline{\eta}\, \underline{\eta}'$) essendo le variabili latenti endogene η a media nulla. Per ottenere E($\underline{Y}$ $\underline{Y'}$), si parte dalla relazione (5)

$\underline{\eta}$=B $\underline{\eta}$+Γ $\underline{\xi}$ +$\underline{\zeta}$ => (I-B)$\underline{\eta}$=Γ $\underline{\xi}$ +$\underline{\zeta}$ => $\underline{\eta}$=(I-B)$^{-1}$ (Γ $\underline{\xi}$ +$\underline{\zeta}$)

e quindi la matrice di varianza-covarianza per le variabili η risulta:

E($\underline{\eta}\, \underline{\eta}'$)= (I-B)$^{-1}$(ΓΦΓ′ + Ψ)(I-B)$^{-1}$'

La matrice delle varianze-covarianze per le variabili Y, ricordando la (2.a) diviene:

E($\underline{Y}$ $\underline{Y}$')= $\Lambda_y[(I-B)^{-1}(\Gamma\Phi\Gamma' + \Psi)(I-B)^{-1\prime}]\Lambda_y' + \Theta_\varepsilon$

La matrice campionaria **S** è formata dalle matrici di varianza-covarianza delle variabili osservate X, delle variabili osservate Y e dalle covarianze tra le variabili osservate X e le osservate Y e quindi si determina l'equivalente covarianza del modello

E($\underline{X}$ $\underline{Y}$')= $\Lambda_x \Phi \Gamma'[(\mathrm{I}-\mathrm{B})^{-1}]'\Lambda_y'$

La matrice **Σ** diviene:

$$\boldsymbol{\Sigma}=\begin{bmatrix} \Lambda_y[(I-B)^{-1}(\Gamma\Phi\Gamma' + \Psi)(I-B)^{-1\prime}]\Lambda_y' + \Theta_\varepsilon & \Lambda_y\Gamma[(\mathrm{I}-\mathrm{B})^{-1}]\Phi\,\Lambda_x' \\ \Lambda_x\Phi\Gamma'[(\mathrm{I}-\mathrm{B})^{-1}]'\Lambda_y' & \Lambda_x\Phi\Lambda_x' + \Theta_\delta \end{bmatrix}$$

2.3 Effetti diretti, indiretti e totali

Le relazioni tra variabili possono essere di tre tipi: l'associazione, l'effetto diretto e quello indiretto. L'associazione è una relazione non direzionale tra 2 variabili e non è un effetto in termini causali. I parametri, che esprimono questa misura, sono le covarianze. L'effetto diretto è una relazione direzionale e causale tra 2 variabili ed è espresso direttamente dai parametri strutturali e dai loadings. L'effetto indiretto è una relazione tra 2 variabili mediata da almeno un'altra variabile. L'effetto indiretto è un effetto causale e non è espresso direttamente da un parametro del modello. La somma dell'effetto diretto e di quello indiretto è l'effetto totale, che è ancora un effetto causale. In questi termini si può ridefinire la definizione di variabili esogene e quella di variabili endogene. Le prime sono quelle che ricevono al più un'influenza non direzionale, le seconde quelle che ricevono quella direzionale. Come si è visto per i modelli causali categoriali, anche nei modelli SEM l'effetto indiretto e quello totale possono essere ricavati dai parametri stimati nel modello. Uno schema è proposto da Bollen (1989):

	Effetti su		
	η	y	x
Effetti di ξ			
Diretti	Γ	0	Λ_x
Indiretti	$(\mathrm{I}-\mathrm{B})^{-1}\Gamma-\Gamma$	$\Lambda_y(\mathrm{I}-\mathrm{B})^{-1}\Gamma$	0
totali	$(\mathrm{I}-\mathrm{B})^{-1}\Gamma$	$\Lambda_y(\mathrm{I}-\mathrm{B})^{-1}\Gamma$	Λ_y
Effetti di η			
Diretti	B	Λ_y	0
Indiretti	$(\mathrm{I}-\mathrm{B})^{-1}-\mathrm{I}-\mathrm{B}$	$\Lambda_y(\mathrm{I}-\mathrm{B})^{-1}-\Lambda_y$	0
totali	$(\mathrm{I}-\mathrm{B})^{-1}-\mathrm{I}$	$\Lambda_y(\mathrm{I}-\mathrm{B})^{-1}$	0

L'effetto totale si può ottenere anche utilizzando lo stesso metodo usato per il modello causale categoriale. Un esempio si trova in Bollen (1989), utilizzato dall'autore per spiegare la causalità e precisamente il concetto di "isolation". Si ipotizza di avere un modello ad equazioni simultanee con tutte le variabili osservate del tipo:

$$Y_{\mathbf{1}} = \gamma_{11}X_1 + \zeta_1$$

$$Y_{\mathbf{2}} = \beta_{21}Y_1 + \gamma_{21}X_1 + \zeta_2$$

In una situazione così, X_1 influisce sia direttamente che indirettamente su $Y_{\mathbf{2}}$ attraverso $Y_{\mathbf{1}}$. Come per il modello causale si può anche osservare solo la relazione tra $Y_{\mathbf{2}}$ e X_1, cioè

$$Y_{\mathbf{2}} = \gamma_{21}{}^{*}X_1 + \zeta_2{}^{*}$$

dove $\zeta_2{}^{*} = \beta_{21}Y_1 + \zeta_2$

Non essendo più in una situazione di equazioni simultanee, il parametro si stima con un OLS cioè:

$$\hat{\gamma}_{21}{}^{*} = \frac{Cov(X_1Y_2)}{Var(X_1)} \qquad (6)$$

Il suo limite in probabilità è

$$plim(\hat{\gamma}_{21}{}^{*}) = plim\left(\frac{Cov(X_1Y_2)}{Var(X_1)}\right) = \frac{Cov(X_1Y_2)}{Var(X_1)} = \gamma_{21} + \beta_{21}\gamma_{11} = \gamma_{21}{}^{*}$$

dimostrando che $\gamma_{21}{}^{*}$ è uguale non solo asintoticamente all'effetto totale, cioè $\gamma_{21} + \beta_{21}\gamma_{11}$.

Con questo esempio si è verificato che come per i modelli loglineari causali completi a 3 variabili A, B e C, l'effetto totale si studia nella tabella di frequenza AC; in un modello SEM completo a 3 variabili X_1, Y_1 e $Y_{\mathbf{2}}$ l'effetto totale si studia dalla matrice di varianza-covarianza $X_1\ Y_2$.
Nel seguito, come si è effettuato nei modelli causali categoriali, anche per i modelli SEM si costruisce dai parametri SEM un dataset composto da 3 variabili (X, Y e Z) e si prova a calcolare l'effetto totale dalla matrice di varianza-covarianza XYZ e da quella XZ. I 5 modelli qui di seguito proposti rispecchiano i modelli causali log-lineari del capitolo 1, come evidenziato nella tabella 1. Come in quel caso, quindi, i parametri stimati ricalcano esattamente quelli del processo generatore dei dati, per cui non è di interesse la parte inferenziale. Sempre per esigenze di confrontabilità, nel seguito si considerano solo variabili osservate.

Tabella 1: equivalenza tra le denominazioni dei modelli nella metodologia SEM e in quella log-lineare causale

modello SEM		modello log-lineare causale	
modello completo		modello completo	
sub modelli:	con solo effetto indiretto	sub modelli:	con solo effetto indiretto
	primo caso solo effetto diretto		Solo effetto diretto
	secondo caso solo effetto diretto		Effetto diretto e cella
modello completo con effetto totale nullo		modello completo con effetto totale nullo	

Nei modelli SEM i dati devono essere espressi in forma di matrice varianza- covarianza e perciò si calcolano i seguenti valori:

$$Var(Y) = {\gamma_y}^2\phi + \psi_y$$

$$Var(Z) = (\beta\gamma_y + \gamma_z)^2\phi + \beta^2\psi_y + \psi_z$$

$$Var(X) = \phi$$

$$Cov(YZ) = \left(\beta\gamma_y + \gamma_z\right)\phi\gamma_y + \beta\psi_y$$

$$Cov(YX) = \gamma_y\phi$$

$$Cov(ZX) = \left(\beta\gamma_y + \gamma_z\right)\phi$$

I valori dei parametri si scelgono in modo tale che la matrice di varianza-covarianza sia definita positiva. Gli errori delle Y sono incorrelati e quindi la matrice ψ è diagonale.
Il dataset di partenza, relativo al modello completo, ha tutti i parametri. I 3 sub-modelli hanno ciascuno un parametro strutturale posto uguale a 0 e rispettivamente: nel modello "con solo effetto indiretto" $\gamma_z = 0$, nel modello "primo caso solo effetto diretto" β=0, nel modello "secondo caso solo effetto diretto" γ_y=0.

2.3.1 Modello completo

In questo modello X influisce direttamente sia su Z che su Y e quest'ultima influisce direttamente su Z, perciò la matrice dei γ è completa e quella dei β è triangolare.

$$\gamma_y = -0.5$$
$$\gamma_z = 1.3$$
$$\beta = 2$$

$$\phi = 7$$
$$\psi_y = 3.12$$
$$\psi_z = 19.9$$

La matrice che ne deriva è

Tabella 2: matrice varianza-covarianza modello completo

	Y	Z	X
Y	4.87		
Z	5.19	33.01	
X	-3.5	2.1	7

La stima Lisrel è

Tabella 3: stima modello completo dalla matrice di varianza-covarianza XYZ

BETA	GAMMA	PHI	PSI
BETA Y Z -------- ------ Y - - - - Z 2.00 - -	GAMMA X -------- Y -0.50 Z 1.3	PHI X -------- 7.00	PSI Note: This matrix is diagonal. Y Z -------- -------- 3.12 19.90
totale	indiretto		
Total Effects X -------- Y -0.50 Z 0.30	Indirect Effects X -------- Y - - Z -1.00		

Tabella 4: stima modello completo dalla matrice di varianza-covarianza XZ

GAMMA	PHI	PSI
X -------- Z 0.30	X -------- 7.00	Z -------- 32.38

Nella tabella 3 e in quella 4 risulta che l'effetto totale della variabile esogena X sulla variabile endogena Z è uguale a 0.3 con entrambe le metodologie [effettto totale= 1.3+ (-0.5)·(2.00)=0.3]

2.3.2 Modello con solo effetto indiretto

In questo modello, la variabile esogena X influisce direttamente solo sulla variabile endogena Y e quest'ultima influisce direttamente sulla endogena Z, quindi la matrice dei γ ha solo γ_y e quella dei β è triangolare.

$$\gamma_y = -0.5$$
$$\gamma_z = 0$$
$$\beta = 2$$
$$\phi = 7$$
$$\psi_y = 3.12$$
$$\psi_z = 19.9$$

La matrice che ne deriva è

Tabella 5: matrice varianza-covarianza modello solo effetto indiretto

	Y	Z	X
Y	4.87		
Z	9.74	39.38	
X	-3.5	-7	7

La stima Lisrel è

Tabella 6: stima modello solo effetto indiretto dalla matrice di varianza-covarianza XYZ

<table>
<tr><td>BETA

Y Z
-------- --------
Y - - - -
Z 2.00 - -</td><td>GAMMA

X

Y -0.50
Z 0.00</td><td>PHI

X

7.00</td><td>PSI
Note: This matrix is diagonal.

Y Z
-------- --------
3.12 19.90</td></tr>
<tr><td colspan="2">totale</td><td colspan="2">indiretto</td></tr>
<tr><td colspan="2">Total Effects

X

Y -0.50
Z -1.00</td><td colspan="2">Indirect Effects

X

Y - -
Z -1.00</td></tr>
</table>

Il parametro γ_z, anche se supposto presente nel modello, è correttamente stimato uguale a 0.

Tabella 7: stima modello con solo effetto indiretto dalla matrice di varianza-covarianza XZ

GAMMA	PHI	PSI
X -------- Z -1.00	X -------- 7.00	Z -------- 32.38

Le tabelle 6 e 7 indicano che l'effetto totale della variabile esogena X sulla variabile endogena Z è uguale a -1.00 con entrambe le procedure [effetto totale = - 0.5· 2=-1.00].

2.3.3 Modello primo caso con solo effetto diretto

In questo modello, la variabile esogena X influisce direttamente sia sulla variabile endogena Y che sulla variabile endogena Z, ma tra Z e Y non esiste alcuna relazione, quindi la matrice dei γ è completa e quella dei β non esiste.

$$\gamma_y = -0.5$$
$$\gamma_z = 1.3$$
$$\beta = 0$$
$$\phi = 7$$
$$\psi_y = 3.12$$
$$\psi_z = 19.9$$

La matrice che ne deriva è

Tabella 8: matrice varianza-covarianza modello primo caso solo effetto indiretto

	Y	Z	X
Y	4.87		
Z	-4.55	31.73	
X	-3.5	9.1	7

La stima Lisrel è

Tabella 9: stima modello primo caso con solo effetto indiretto dalla matrice di varianza-covarianza XYZ

<table>
<tr><td>BETA

Y Z
-------- --------
Y - - - -
Z 0.00 - -</td><td>GAMMA

X

Y -0.50
Z 1.30</td><td>PHI

X

7.00</td><td>PSI
Note: This matrix is diagonal.

Y Z
-------- --------
3.12 19.90</td></tr>
<tr><td colspan="2">totale</td><td colspan="2">indiretto</td></tr>
<tr><td colspan="2">Total Effects

X

Y -0.50
Z 1.30</td><td colspan="2"></td></tr>
</table>

Il parametro β, anche se supposto presente nel modello, è esattamente stimato uguale a 0.

Tabella 10: stima modello primo caso con solo effetto diretto dalla matrice di varianza-covarianza XZ

<table>
<tr><td>GAMMA

X

Z 1.30</td><td>PHI

X

7.00</td><td>PSI

Z

19.90</td></tr>
</table>

Nelle tabelle 9 e 10 è evidenziato che l'effetto totale della variabile esogena X sulla endogena Z è uguale a 1.3 con entrambe le procedure.

2.3.4 Modello secondo caso con solo effetto diretto

Questo modello può ancora essere considerato un sotto-modello di quello completo ($\gamma_y = 0$), sebbene la variabile Y divenga nella sostanza da endogena a esogena, poiché la variabile esogena X influisce solo sulla variabile endogena Z, su cui influisce anche Y. Tuttavia, è ugualmente possibile mantenere per coerenza la notazione utilizzata finora.

$$\gamma_y = 0$$
$$\gamma_z = 1.3$$
$$\beta = 2$$
$$\phi = 7$$
$$\psi_y = 3.12$$
$$\psi_z = 19.9$$

La matrice che ne deriva è:

Tabella 11: matrice varianza-covarianza modello secondo caso solo effetto diretto

	Y	Z	X
Y	3.12		
Z	6.24	44.21	
X	0	9.1	7

Tabella 12: stima modello primo caso con solo effetto diretto dalla matrice di varianza-covarianza XYZ (inserendo nel processo di stima l'ipotesi γ_y)

<table>
<tr>
<td>BETA

Y Z
-------- --------
Y - - - -
Z 2.00 - -</td>
<td>GAMMA

X

Y - -
Z 1.30</td>
<td>PHI

X

7.00</td>
<td>PSI
Note: This matrix is diagonal.

Y Z
-------- --------
3.12 19.90</td>
</tr>
<tr>
<td colspan="2">totale</td>
<td colspan="2">indiretto</td>
</tr>
<tr>
<td colspan="2">Total Effects

X

Y - -
Z 1.30</td>
<td colspan="2"></td>
</tr>
</table>

Tabella 13: stima modello secondo caso con solo effetto diretto dalla matrice di varianza-covarianza XZ

<table>
<tr>
<td>GAMMA

X

Z 1.30</td>
<td>PHI

X

7.00</td>
<td>PSI

Z

32.38</td>
</tr>
</table>

Nella tabelle 12 e 13 risulta che l'effetto totale della variabile esogena X sulla variabile endogena Z è uguale a 1.3 con entrambe le procedure (stima dalla matrice di varianza-covarianza XYZ e stima dalla matrice di varianza-covarianza XZ).

2.3.5 Modello completo con effetto totale nullo

In questo modello, la variabile esogena X influisce direttamente su entrambe le variabili endogene Y e Z, Y influisce direttamente su Z perciò la matrice dei γ e quella dei β sono complete. Per rendere nullo l'effetto totale della variabile esogena X sulla variabile endogena Z si cambia il valore del parametro γ_z rispetto a quello usato nel dataset della tabella 1. Precisamente si è posto il parametro strutturale γ_z uguale all'opposto dell'effetto indiretto, rappresentato dal prodotto tra l'effetto diretto X → Y e l'effetto diretto Y → Z (effetto indiretto = $\gamma_y \cdot \beta$ = -0.5 · 2=1)

$$\gamma_y = -0.5$$
$$\gamma_z = 1$$
$$\beta = 2$$
$$\phi = 7$$
$$\psi_y = 3.12$$
$$\psi_z = 19.9$$

La matrice che ne deriva è

Tabella 14: matrice varianza-covarianza modello con effetto totale nullo

	Y	Z	X
Y	4.87		
Z	6.24	32.38	
X	-3.5	0	7

La stima Lisrel è

Tabella 15: stima modello con effetto totale nullo dalla matrice di varianza-covarianza XYZ

BETA	GAMMA	PHI	PSI
Y Z -------- -------- Y - - - - Z 2.00 - -	X -------- Y -0.50 Z 1.00	X -------- 7.00	Note: This matrix is diagonal. Y Z -------- -------- 3.12 19.90
totale		indiretto	
Total Effects X -------- Y -0.50 Z **0**		Indirect Effects X -------- Y - - Z -1.00	

L'effetto totale è stato correttamente identificato uguale a 0. Solo per problemi riguardanti i "comandi" del programma Lisrel non è possibile stimare un effetto totale nullo dalla matrice di varianza-covarianza ZX in quanto richiede o di inserire l'esogeneità delle variabili o l'equivalente vincolo uguale a 0, che rende esogena la variabile. Questo problema si presenta nel caso della stima dell'effetto totale nullo e anche stimando "il modello secondo caso solo effetto diretto " (dataset tabella 11) senza imporre che la variabile Y sia esogena o che $\gamma_y$sia uguale a 0.

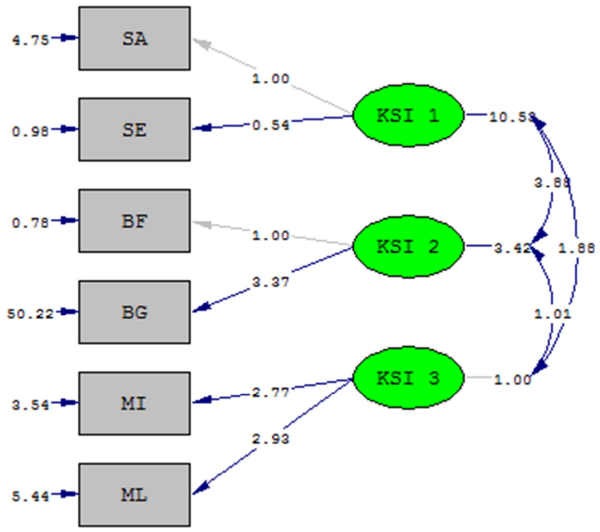
Tabella 16: stima modello secondo caso con solo effetto diretto dalla matrice di varianza-covarianza XYZ (senza imporre Y esogena o γ_y=0)

BETA	GAMMA	PHI	PSI
Y Z -------- -------- Y - - - - Z 2.00 - -	X -------- Y 0.50 Z 1.30	X -------- 7.00	Note: This matrix is diagonal. Y Z -------- -------- 3.12 19.90
totale		indiretto	
Total Effects X -------- Y 0.50 Z 2.30		Indirect Effects X -------- Y - - Z 1.00	

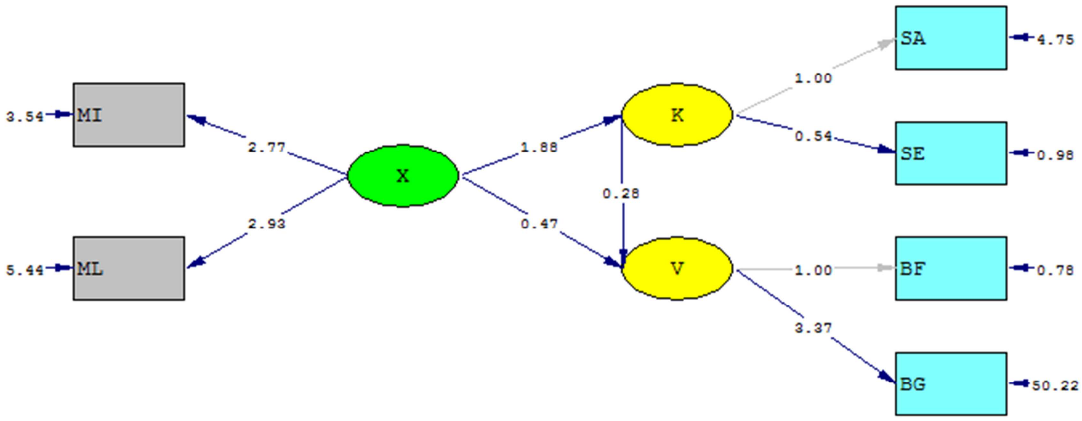
Il parametro γ_y è erroneamente identificato uguale a 0.5. Il parametro γ_y è erroneamente identificato uguale a 0.5. La richiesta di Lisrel di esogeneità di una variabile non è un vero problema in quanto, osservando la matrice di varianza-covarianza XYZ (tabella 11) con Cov(YX) uguale 0, risulta Y esogena (γ_y=0).

2.4 Path Diagram

Il path diagram è un grafico che spiega le relazioni tra le variabili. Le variabili osservate sono rappresentate da un rettangolo, quelle latenti da una ellisse. I colori, usati da Lisrel, distinguono se le variabili latenti sono esogene o endogene: il verde per le esogene (le ξ) e il giallo per le endogene (le η). Le variabili osservate cambiano colore se esse sono indicatori di endogene o di esogene: le X sono grigie, le Y azzurre. Le relazioni causali sono rappresentate da una freccia unidirezionale, le correlazioni da una curva bi-direzionale. In generale entrambe sono blu, ma se fissate divengono grigie. Ecco due esempi di un modello che sarà analizzato nel capitolo successivo: il primo si riferisce ad una analisi CFA, il secondo a un modello strutturale a variabili latenti.



Il grafico rappresenta un modello CFA e quindi le variabili osservate sono tutte trattate come X e quelle latenti come ξ, essendo tutte esogene, cioè senza relazioni causali tra di loro. Le X sono rappresentate da rettangoli grigi, le ξ da ellissi verdi. I loadings sono rappresentati dalle frecce che vanno dalle ksi alle X: blu per quelli stimati, grigie per quelli fissati (in questo caso per porre la scala). Le frecce unidirezionali blu (quelle con i valori di 4.75, 0.98, ecc) che vanno verso le X sono gli errori δ. Le curve bidirezionali (per esempio quella con valore 1.01) che collegano le ksi sono le loro covarianze, mentre i valori 3.42, 1 e 10.53 sono le loro varianze.



Questo è un modello strutturale e quindi c'è la distinzione fra esogeno ed endogeno. L'unica latente esogena è rappresentata dall'ellisse verde e le 2 latenti endogene dall'elisse gialla. Gli indicatori, cioè le variabili osservate, delle esogene sono rappresentati da rettangoli grigi, quelli delle endogene da rettangoli azzurri. Le frecce che collegano le latenti ai loro indicatori sono i loadings, le frecce che collegano tra loro le endogene sono i parametri strutturali, cioè i β e i γ. Le frecce a cui è associato un valore e che sono direzionate verso le osservate (per esempio quella associata a 5.44 e quella associata a 50.22) sono le varianze degli errori.

Capitolo 3:
SEM vs Latent Class Analysis

3.1 Introduzione

Nel capitolo 1 si è ipotizzato che le variabili categoriali utilizzate nell'analisi log-lineare siano direttamente osservate. Analogamente da quanto visto nel capitolo 2 per le varibili continue, esistono, però, variabili che non possono essere osservate direttamente: il solo modo per ottenerle è quello di cercare variabili che possano essere osservate direttamente e che contengano informazioni teoriche su queste. A seconda della tipologia del binomio osservate/latenti si utilizza un diverso approccio. Bartholomew (1998) lo sintetizza nella seguente tabella:

Variabili Latenti	Variabili Manifeste	
	Continue	Categoriali
Continue	Analisi fattoriale	Latent trait analysis
Categoriali	Latent profile analysis	Latent Class Analysis

L'analisi fattoriale, che sta alla base dei modelli SEM, è stata analizzata nel capitolo precedente, mentre i metodi latent trait analysis e latent profile analysis non sono oggetto di questa tesi. Per quanto riguarda la Latent Class Analysis (LCA), le variabili categoriali possono essere intervallari, ordinali o nominali e quindi si ha una ulteriore suddivisione del metodo LCA ben sintetizzata dalla tabella proposta da Heinen (1996):

Variabili Latenti	Variabili Manifeste discrete		
	Nominali	Ordinali	Intervallari
Nominali	LCA		LCA per rating data (Rost)
Ordinali		LCA con classi ordinate (Croon)	
Intervallari	1)LCA con restrizioni lineari (Haberman) 2)model response nominal (Bock)	Graded response model (Samejiama)	1)LCA con linear-by-linear restrictions (Haberman) 2)partial credit model (Masters) 3)rating scale model (Andrich)

La natura del dataset (cioè se formato da dati categoriali o continui) dovrebbe portare a scegliere se utilizzare un modello SEM o un modello loglineare causale a variabili latenti per analizzare i dati. Tuttavia, in gran parte degli ambiti applicativi si preferisce usare il metodo SEM anziché quello causale, anche se il dataset è formato da variabili ordinali, spesso considerando valida l'approssimazione normale in presenza di almeno 7 categorie. L'analisi di entrambi i modelli (SEM e causale con dati categoriali) può essere suddivisa in una parte fattoriale e in una parte strutturale. In questo capitolo vengono paragonati i due metodi (SEM e modello loglineare causale a variabili latenti).

3.2 Modelli a classi latenti

L'assunzione base della LCA è l'indipendenza locale, che consiste nell'ipotizzare che le variabili osservate (ad esempio A e B) siano dipendenti solo attraverso la loro relazione con le variabili latenti (in questo caso θ), cioè

$A \perp B|\theta$

Questa assunzione è indebolita da Hagenaars (1988,1993), il quale ha permesso la correlazione tra variabili osservate. Esistono 2 tipi di parametrizzazione della LCA, uno in termini di probabilità condizionate, l'altro in termini di parametri log-lineari (Heinen 1996); si ipotizzi di avere 3 variabili osservate (A ,B e C) ed una latente (θ):

- $\pi^{A=i\ B=j\ C=k\ \theta=t} = \pi^{\theta=t}\pi^{A=iB=jC=k|\ \theta=t} = \pi^{\theta=t}\pi^{A=i\ |\theta=t}\pi^{B=j\ |\theta=t}\pi^{C=k\ |\ \theta=t}$

dove

$\pi^{\theta=t}$= probabilità di essere nella classe θ= t

$\pi^{A=i\ |\theta=t}$= probabilità condizionata di A= i dato θ= t

- Ln $\pi^{A=iB=jC=k\theta=t} = u + \lambda^{A=i} + \lambda^{B=j} + \lambda^{C=k} + \lambda^{\theta=t} + \lambda^{A=i\theta=t} + \lambda^{B=j\theta=t} + \lambda^{C=k\theta=t}$

Si possono naturalmente esprimere le probabilità condizionate in termini log-lineari (Heinen 1996):

$$\pi^{A=i|\theta=t} = \frac{\pi^{A=i\ \theta=t}}{\pi^{\theta=t}} = \frac{\sum_j \sum_k \pi^{A=i\ B=j\ C=k\ \theta=t}}{\sum_i \sum_j \sum_k \pi^{A=iB=jC=k\theta=t}}$$
$$= \frac{exp(\lambda^{A=i} + \lambda^{A=i\theta=t}) \sum_j \sum_k exp(u + \lambda^{B=j} + \lambda^{C=k} + \lambda^{\theta=t} + \lambda^{B=j\ \theta=t} + \lambda^{C=k\ \theta=t})}{\left[\sum_i exp(\lambda^{A=i} + \lambda^{A=i\ \theta=t})\right] \cdot \left[\sum_j \sum_k exp(u + \lambda^{B=j} + \lambda^{C=k} + \lambda^{\theta=t} + \lambda^{B=j\ \theta=t} + \lambda^{C=k\ \theta=t})\right]}$$
$$= \frac{exp(\lambda^{A=i} + \lambda^{A=i\ \theta=t})}{\sum_i exp(\lambda^{A=i} + \lambda^{A=i\ \theta=t})}$$

Un metodo utilizzato per stimare modelli con classi latenti è l'"EM", dove E sta per valore atteso (in inglese expectation) e M per massimizzazione. La procedura base dell' "EM" (Haberman 1979, Hagenaars 1990, Croon et al. 2009) si sviluppa nel seguente modo. E' una stima che procede a 2 passi: nello step E si calcola il valore atteso della probabilità congiunta delle variabili osservate e latenti; nello step M le probabilità congiunte sono considerate come date e le stime ML delle probabilità sono calcolate utilizzando il metodo per le variabili osservate. I 2 steps continuano fino ad arrivare ad una convergenza, che si può avere quando le successive stime non cambiano di molto.

1) si sceglie un valore iniziale $\hat{\pi}^{ABC\theta\,(0)}$
2) step "E": si calcola $\hat{p}^{ABC\theta\,(s)} = p^{ABC}\frac{\hat{\pi}^{ABC\theta\,(s-1)}}{\hat{\pi}^{ABC\theta}} = p^{ABC}\hat{\pi}^{\theta|ABC(s)}$
   dove p^{ABC} rappresenta la proporzione osservata nel dataset, $\hat{\pi}^{ABC\theta\,(s-1)}$ con s=1 è il valore iniziale e con s>1 è il valore trovato nello step "M" precedente
3) step "M": si calcola $\hat{\pi}^{ABC\theta\,(s+1)} = \hat{p}^{A|\theta\,(s)}\hat{p}^{B|\theta\,(s)}\hat{p}^{C|\theta\,(s)}\hat{p}^{\theta\,(s)}$
   dove le probabilità condizionate $\hat{p}^{A|\theta\,(s)}, \hat{p}^{B|\theta\,(s)}, \hat{p}^{C|\theta\,(s)} e\ \hat{p}^{\theta\,(s)}$ sono calcolate come funzione di $\hat{p}^{ABC\theta\,(s)}$dello step "M":

$$\hat{p}^{\theta\,(s)} = \sum_i\sum_j\sum_k \hat{p}^{ABC\theta\,(s)}$$
$$\hat{p}^{A|\theta\,(s)} = \sum_i\sum_k \left(\hat{p}^{ABC\theta\,(s)}/\hat{p}^{\theta\,(s)}\right)$$
$$\hat{p}^{B|\theta\,(s)} = \sum_i\sum_k \left(\hat{p}^{ABC\theta\,(s)}/\hat{p}^{\theta\,(s)}\right)$$
$$\hat{p}^{C|\theta\,(s)} = \sum_i\sum_j \left(\hat{p}^{ABC\theta\,(s)}/\hat{p}^{\theta\,(s)}\right)$$

3.2.1 MODELLO CAUSALE CON VARIABILI LATENTI: una trasposizione del SEM nel categoriale

Le variabili non osservate possono essere in numero maggiore di 1, questo comporta un primo passo verso il modello causale con variabili latenti: A e B sono indicatori di θ, C e D sono indicatori di ω

$$\pi^{A=i\ B=j\ C=k\ D=l\ \theta=t\ \ \omega=w} = \pi^{\theta=t\ \omega=w}\pi^{A=i|\theta=t}\pi^{B=j|\theta=t}\pi^{C=j|\omega=w}\pi^{D=l|\omega=w}$$

In un modello tradizionale si presuppone una relazione bidirezionale tra θ e ω, in un modello causale si presuppone una relazione gerarchica fra le 2 variabili latenti:

modello tradizionale ω⟷θ =>$\pi^{\theta=t\omega=w}$

modello causale ω⟶θ =>$\pi^{\theta=t\omega=w}$=$\pi^{\theta=t|\omega=w}\pi^{\omega=w}$

Come visto per il modello causale senza variabili latenti, la probabilità congiunta ABCDθω può essere decomposta e, tenendo in considerazione l'ordine gerarchico, le varie probabilità possono essere stimate separatamente . Se i dati sono clustered (vedi Capitolo 1) una stima separata delle singole parti non è permessa e quindi tutte le singole parti del modello devono essere stimate simultaneamente. La probabilità congiunta $\pi^{A=i\,B=j\,C=k\,D=l\,\theta=t\omega=w}$è suddivisa in una parte di misura (derivata dall'analisi fattoriale), $\pi^{A=i|\theta=t}\pi^{B=j|\theta=t}\pi^{C=j|\omega=w}\pi^{D=l|\omega=w}$,e in una strutturale $\pi^{\theta=t\omega=w}$.

3.3 Analisi fattoriale

L'analisi fattoriale, così chiamata da Vermunt e Magidson (2001) anche nei modelli log-lineari per similitudine con i SEM, è la prima fase di un'analisi dei dati e serve a diminuire il numero delle variabili, presupponendo che il legame tra queste sia dovuto a delle variabili latenti . Le variabili osservate vengono definite indicatori delle variabili non osservate. Con variabili continue la metodologia più utilizzata è l'analisi fattoriale di 2 tipi: la confermativa (CFA) e l'esplorativa(EFA). Nel seguito si definisce un parallelo tra queste 2 tecniche e le possibili analisi con variabili categoriali.

3.3.1 Analisi fattoriale esplorativa (EFA)

Nel'EFA si cerca di raggruppare le X (v. osservate) in ξ (v. non osservate) determinando quante ξ stanno sotto ai dati osservati; in un modello LCA si cerca di raggruppare le X (v. osservate) in ξ (v. non osservate) o ricercando il numero di ξ o determinando quante categorie ha la supposta unica ξ. Per sintetizzare, l'EFA, in un modello SEM, ha un unico modo di procedere, in un LCA ne ha 2. E' stato dimostrato da Vermunt che, per esempio, un modello con una variabile latente a 4 categorie può essere ricondotto a un modello con 2 variabili latenti dicotomiche e correlate.
Qui di seguito si mostra come viene esplicitata un'analisi esplorativa. Il dataset che si utilizza per paragonare i 2 metodi (SEM e modello LCA) è composto da 6 variabili osservate a media nulla (A,B,C,D,E ed F) che sono raggruppabili in 3 latenti (V,Z,K) anch'esse a media nulla.

SEM: EFA

$$\begin{matrix} A \\ B \\ C \\ D \\ E \\ F \end{matrix} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \\ \lambda_{41} & \lambda_{42} & \lambda_{43} \\ \lambda_{51} & \lambda_{52} & \lambda_{53} \\ \lambda_{61} & \lambda_{62} & \lambda_{63} \end{bmatrix} \begin{bmatrix} V \\ Z \\ K \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_5 \\ \delta_6 \end{bmatrix} \qquad (1)$$

La formula (1) mostra che le variabili osservate (A,B,C,D,E ed F) sono tutte indicatrici delle 3 variabili latenti, come richiede l'EFA. I parametri λ sono i loadings. I δ sono termini di errore a media zero e incorrelati fra loro (la loro matrice di varianza – covarianza, TD, è diagonale). Le variabili latenti (V,Z,K) sono a media nulla e in partenza assunte incorrelate (la loro matrice di varianza-covarianza,PHI, è diagonale). I modelli con PHI non diagonale richiedono alcune restrizioni di identificazione ma portano al medesimo modello dal punto di vista dell'adattamento ai dati

$$PHI=\begin{bmatrix}\phi_{11} & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \phi_{33}\end{bmatrix}$$

$$TD=\begin{bmatrix}\theta_{11} & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \theta_{66}\end{bmatrix}$$

LCA: EFA

$$\pi^{A=a\ B=b\ C=c\ D=d\ E=e\ F=f\ V=v\ Z=z\ K=k} = \pi^{V=v}\pi^{Z=z}\pi^{K=k}\pi^{A=a|V=v\ Z=z\ K=k}\pi^{B=b|V=v\ Z=z\ K=k}\pi^{C=c|V=v\ Z=z\ K=k}\pi^{D=d|V=v\ Z=z\ K=k}\pi^{E=e|V=v\ Z=z\ K=k}\pi^{F=f|V=v\ Z=z\ K=k} \quad (2)$$

La probabilità congiunta $\pi^{A=a\ B=b\ C=c\ D=d\ E=e\ F=f\ V=v\ Z=z\ K=k}$ è scomposta nella probabilità congiunta delle latenti ($\pi^{V=v\ Z=z\ K=k}$) e nelle osservate condizionate alle latenti ($\pi^{A=a|V=v\ Z=z\ K=k}$, $\pi^{B=b|V=v\ Z=z\ K=k}$, $\pi^{C=c|V=v\ Z=z\ K=k}$, $\pi^{D=d|V=v\ Z=z\ K=k}$, $\pi^{E=e|V=v\ Z=z\ K=k}$, $\pi^{F=f|V=v\ Z=z\ K=k}$).

La congiunta $\pi^{V=v\ Z=z\ K=k}$, grazie all'indipendenza tra V, Z e K , è uguale al prodotto delle 3 marginali ($\pi^{V=v}\pi^{Z=z}\pi^{K=k}$). La (2) può essere scritta anche nella formulazione log-lineare rappresentata nella (3)

$$\begin{cases}\log(\pi^{V=v}) = u^{V} + \lambda^{V=v}\\ \log(\pi^{Z=z}) = u^{Z} + \lambda^{Z=z}\\ \log(\pi^{K=k}) = u^{K} + \lambda^{K=k}\\ \log(\pi^{A=a|V=v\ Z=z\ K=k}) = u^{A|V=v\ Z=z\ K=k} + \lambda^{A=a} + \lambda^{A=a\ V=v} + \lambda^{A=a\ Z=z} + \lambda^{A=a\ K=k}\\ \log(\pi^{B=b|V=v\ Z=z\ K=k}) = u^{B|V=v\ Z=z\ K=k} + \lambda^{B=b} + \lambda^{B=b\ V=v} + \lambda^{B=b\ Z=z} + \lambda^{B=b\ K=k}\\ \log(\pi^{C=c|V=v\ Z=z\ K=k}) = u^{C|V=v\ Z=z\ K=k} + \lambda^{C=c} + \lambda^{C=c\ V=v} + \lambda^{C=c\ Z=z} + \lambda^{C=c\ K=k}\\ \log(\pi^{D=d|V=v\ Z=z\ K=k}) = u^{D|V=v\ Z=z\ K=k} + \lambda^{D=d} + \lambda^{D=d\ V=v} + \lambda^{D=d\ Z=z} + \lambda^{D=d\ K=k}\\ \log(\pi^{E=e|V=v\ Z=z\ K=k}) = u^{E|V=v\ Z=z\ K=k} + \lambda^{E=e} + \lambda^{E=e\ V=v} + \lambda^{E=e\ Z=z} + \lambda^{E=e\ K=k}\\ \log(\pi^{F=f|V=v\ Z=z\ K=k}) = u^{F|V=v\ Z=z\ K=k} + \lambda^{F=f} + \lambda^{F=f\ V=v} + \lambda^{F=f\ Z=z} + \lambda^{F=f\ K=k}\end{cases} \quad (3)$$

I parametri a 2 variabili (per esempio $\lambda^{A=a\ V=v}$) sono l'equivalente dei loadings (cioè del parametro λ_{11} della (1)).

In questo paragrafo si spiega la metodologia usata in un'analisi esplorativa per classi latenti (LCFA) (Vermunt e Magidson 2001). Per farlo si creeranno dei dataset con variabili latenti. Si generano 5 variabili (A B, C, D ed E) la cui dipendenza è dovuta ad A che influisce su tutte. Si utilizza l'effect code e i parametri moltiplicativi sono:

$$\tau^{A} = 0.25$$

$$\tau^{B} = 0.8$$

$$\tau^{C} = 0.2$$

$$\tau^{D} = 4$$

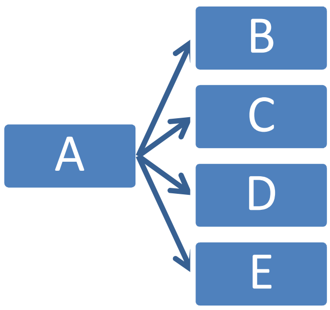
$$\pi^{E} = 0.15$$

$$\tau^{AB} = 0.5$$
$$\tau^{AC} = 2$$

$$\tau^{DA} = 0.8$$

$$\tau^{EA} = 0.25$$

La struttura dei dati è sintetizzata da questo grafico



Le distribuzioni sono:

Tabella 1: B, C, D ed E indicatori di A

a 1 0.0588 2 0.9412	c|a 1|1 0.1379 2|1 0.8621 1|2 0.0099 2|2 0.9901	b|a 1|1 0.1379 2|1 0.8621 1|2 0.7191 2|2 0.2809
d|a 1|1 0.9110 2|1 0.0890 1|2 0.9615 2|2 0.0385	e|a 1|1 0.0014 2|1 0.9986 1|2 0.2647 2|2 0.7353	

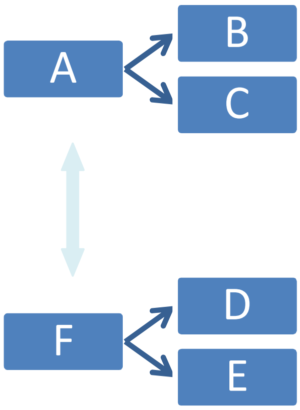
Ipotizzando di non conoscere la distribuzione di A e quindi mettendo come dati la distribuzione congiunta BCDE, la stima del modello risulta:

Tabella 2:stima del modello B,C,D ed E indicatori di A

```
*** LOG-LINEAR PARAMETERS ***

* TABLE A [or P(A)] *

  effect           lambda   exp(lambda)
  A
   1              1.3860    3.9988
   2             -1.3860    0.2501

* TABLE AB [or P(B|A)] *

  effect           lambda   exp(lambda)
  B
   1             -0.2228    0.8002
   2              0.2228    1.2496
  AB
   1 1            0.6929    1.9994
   1 2           -0.6929    0.5001
   2 1           -0.6929    0.5001
   2 2            0.6929    1.9994

* TABLE AC [or P(C|A)] *

  effect           lambda   exp(lambda)
  C
   1             -1.6096    0.2000
   2              1.6096    5.0007
  AC
   1 1           -0.6931    0.5000
   1 2            0.6931    2.0000
   2 1            0.6931    2.0000
   2 2           -0.6931    0.5000

* TABLE AD [or P(D|A)] *

  effect           lambda   exp(lambda)
  D
   1              1.3863    4.0002
   2             -1.3863    0.2500
  AD
   1 1            0.2231    1.2500
   1 2           -0.2231    0.8000
   2 1           -0.2231    0.8000
   2 2            0.2231    1.2500

* TABLE AE [or P(E|A)] *

  effect           lambda   exp(lambda)
  E
   1             -1.8832    0.1521
   2              1.8832    6.5744
  AE
   1 1            1.3724    3.9447
   1 2           -1.3724    0.2535
   2 1           -1.3724    0.2535
   2 2            1.3724    3.9447
```

```
*** LATENT CLASS OUTPUT ***

          A  1    A  2
        0.9411  0.0589
  C  1  0.0099  0.1379
  C  2  0.9901  0.8621
  B  1  0.7191  0.1381
  B  2  0.2809  0.8619
  D  1  0.9615  0.9110
  D  2  0.0385  0.0890
  E  1  0.2647  0.0015
  E  2  0.7353  0.9985
```

Nella colonna 2 della tabella (LATENT CLASS OUTPUT) si vedono le probabilità condizionate degli indicatori rispetto alle latenti, per esempio 0.9901 è la P(C=2|A=1), cioè la probabilità condizionata dell'indicatore C data la latente A=1. In un'analisi fattoriale in cui si ipotizza un solo fattore i valori del "LATENT CLASS OUTPUT"non aggiungono alcuna informazione. Essi divengono importanti in un'analisi fattoriale a 2 o più fattori perché possono aiutare a capire a quale indicatore associare la latente. L'analisi fattoriale esplorativa serve per esplorare le relazioni tra le variabili osservate (cioè quelle del dataset) e determinare quando le relazioni possono essere spiegate dalla presenza di variabili latenti ed è stata introdotta nei modelli a classi latenti da Vermunt. Per mostrare l'utilizzo del "LATENT CLASS OUTPUT"nel caso a 2 fattori, si simula un modello con 4 indicatori (BCDE) e 2 latenti (A e F), così suddivisi:



dove la relazione tra B e C è dovuta ad A e quella tra D ed E è dovuta ad F. Per fare un dataset su cui applicare un'analisi fattoriale si ipotizza che le variabili B, C, D ed E siano osservate e che A e F siano latenti. Per creare il dataset si calcolano le condizionate B|A, C|A, D|F e E|F, così da ottenere la condizionata BCDE|AF (cioè P(BCDE|AF)=P(B|A)P(C|A)P(D|F)P(E|F)). Ipotizzando A correlato a F, P(AF) è diverso da P(A)P(F). Il dataset con tutte le 6 variabili osservate è costruito moltiplicando P(BCDE|AF) per P(AF). Per ipotizzare A e F latenti, si calcola la congiunta P(BCDE) che è il dataset utilizzato nella stima.

Tabella 3: B e C indicatori di A, D ed E indicatori di F, A e F latenti correlate

fa		b|a		d|f	
11	0.5503	11	0.6653	11	0.1042
12	0.0015	12	0.0041	12	0.9264
21	0.2936	21	0.3347	21	0.8958
22	0.1546	22	0.9959	22	0.0736
		c|a		e|f	
		11	0.5737	11	0.0004
		12	0.0187	12	0.5711

	21	0.4263	21	0.9996
	22	0.9813	22	0.4289

In un'analisi fattoriale, in generale, il problema non è solo determinare il binomio indicatore-fattore, ma anche determinare il numero di fattori. Per scegliere il numero di fattori si stima il modello a 1 fattore, poi quello a 2 e si confronta il BIC: se il BIC a 1 fattore è minore di quello a 2, l'analisi è conclusa e si opta per 1 fattore; se il BIC a 2 fattori è minore di quello a 1 si prosegue stimando il modello a 3 e si confronta nuovamente il BIC. La ricerca del modello migliore si continua finché il BIC del modello ad H+1 fattori non è maggiore del BIC del modello ad H fattori e si opta per quello ad H. Si stima il modello ipotizzando i fattori latenti incorrelati . Stimando i dati della tabella 6 con un modello a 1 fattore, il BIC risulta 4948; con quello a 2 fattori il BIC risulta 4937. Dall'analisi del BIC sembra migliore il modello a 2 fattori. Si prova a stimare un modello a 3 fattori e risultando il BIC 4972 si sceglie il modello a 2 fattori. Bartholomew, invece di usare il BIC, controlla se i modelli (in questo caso ad 1 e a 2 fattori) si adattano bene ai dati: sceglie il modello più semplice (con minor numero di fattori) che si adatta meglio ai dati. Il modello ad 1 fattore ha un χ^2 di 44.90 con un df di 1 , perciò non è un buon modello (p-value=0). Il modello a 2 fattori ha un χ^2 di 0.0012 con un df di 6, perciò è un buon modello (p-value=0.97). Come col metodo BIC si sceglie il modello con 2 fattori.
Nel processo di stima si utilizza un modello in cui le variabili osservate (B,C,D ed E) sono indicatori di entrambe le variabili latenti (A e F) ipotizzate indipendenti (Magidson, Vermunt 2001) nel modello di stima. I valori del "LATENT CLASS OUTPUT" sono rappresentati nella tabella 4.

Tabella 4:"Latent class output" risultante dalla stima del dataset "B e C indicatori di A; D ed E indicatori di F; A e F latenti correlate"

```
*** LATENT CLASS OUTPUT ***

          F  1    F  1    F  2    F  2
          A  1    A  2    A  1    A  2
         0.0610  0.5567  0.0377  0.3446
  E  1   0.3136  0.0025  0.9959  0.5746
  E  2   0.6864  0.9975  0.0041  0.4254
  D  1   0.8204  0.1171  0.9978  0.9281
  D  2   0.1796  0.8829  0.0022  0.0719
  B  1   0.9875  0.6601  0.9425  0.2870
  B  2   0.0125  0.3399  0.0575  0.7130
  C  1   0.9802  0.5679  0.9149  0.2223
  C  2   0.0198  0.4321  0.0851  0.7777

E = 0.1281,  lambda = 0.7110
```

La probabilità P(E|A=1,F=1) è molto simile alla P(E|A=2,F=1) e questo presuppone che sia corretto porre P(E|A,F)=P(E|F), cioè che E dipende da F, come è effettivamente . La probabilità P(D|A=2,F=2) è molto simile alla P(D|A=2,F=2) e questo presuppone che sia corretto porre P(D|A,F)=P(D|F), cioè che D dipende da F, come è effettivamente . La probabilità P(B|A=1,F=1) è molto simile alla P(B|A=1,F=2) e questo presuppone che sia corretto dire P(B|A,F)=P(B|A), cioè che B dipende da A, come è effettivamente. La probabilità P(C|A=1,F=1) è molto simile alla P(C|A=1,F=2) e questo presuppone che sia corretto dire P(C|A,F)=P(C|A), cioè che C dipende da A, come è effettivamente . Dal "LATENT CLASS OUTPUT" si deduce di associare agli indicatori B e C la latente A, agli indicatori D ed E la latente F.

3.3.2 Analisi fattoriale confermativa (CFA)

La CFA differisce dalla EFA in quanto ipotizza a priori il numero di latenti e il binomio latente-indicatore: in essa si verifica il modello teorico che si ritiene corretto. Il dataset che si utilizza per paragonare i 2 metodi (SEM e modello loglineare causale a variabili latenti) è ancora quello composto da 6 variabili osservate a media nulla (A,B,C,D,E ed F) che sono raggruppabili in 3 latenti (V,Z,K) anch'esse a media nulla, ma il modello teorico presuppone, per esempio, che A e B siano indicatori di V, C e D di Z, E ed F di K e che V,Z e K siano correlate.

SEM: CFA

$$\begin{matrix} A \\ B \\ C \\ D \\ E \\ F \end{matrix} = \begin{bmatrix} \lambda_{11} & 0 & 0 \\ \lambda_{21} & 0 & 0 \\ 0 & \lambda_{32} & 0 \\ 0 & \lambda_{42} & 0 \\ 0 & 0 & \lambda_{53} \\ 0 & 0 & \lambda_{63} \end{bmatrix} \begin{bmatrix} V \\ Z \\ K \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_5 \\ \delta_6 \end{bmatrix} \tag{4}$$

La formula (4) mostra che le variabili osservate sono a 2 a 2 indicatrici delle 3 variabili latenti, come vuole il modello teorico che si vuole testare nel CFA. I parametri λ sono i loadings. I δ sono termini di errore a media zero e incorrelati fra loro (la loro matrice di varianza –covarianza, TD, è diagonale). Le variabili latenti (V,Z,K) sono a media nulla e correlate (la loro matrice di varianza-covarianza,PHI, è completa e simmetrica).

$$PHI = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{12} & \phi_{22} & \phi_{23} \\ \phi_{13} & \phi_{23} & \phi_{33} \end{bmatrix}$$

$$TD = \begin{bmatrix} \theta_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \theta_{66} \end{bmatrix}$$

LCA: CFA

$$\pi^{A=a\ B=b\ C=c\ D=d\ E=e\ F=f\ V=v\ Z=z\ K=k} = \pi^{V=v\ Z=z\ K=K} \pi^{A=a|V=v} \pi^{B=b|V=v} \pi^{C=c|Z=z} \pi^{D=d|Z=z} \pi^{E=e|\ K=k} \pi^{F=f|\ K=k} \tag{5}$$

La probabilità congiunta $\pi^{A=a\ B=b\ C=c\ D=d\ E=e\ F=f\ V=v\ Z=z\ K=k}$ è nuovamente scomposta nella probabilità congiunta delle latenti ($\pi^{V=v\ Z=z\ K=k}$) e nelle condizionate delle osservate ($\pi^{A=a|V=v}$, $\pi^{B=b|V=v}$, $\pi^{C=c|Z=z}$, $\pi^{D=d|Z=z}$, $\pi^{E=e|\ K=k}$, $\pi^{F=f|\ K=k}$), ma queste identificano direttamente i legami fra latenti ed osservate.

La congiunta $\pi^{V=v\ Z=z\ K=k}$, non essendoci l'indipendenza tra V, Z e K , non è uguale al prodotto delle 3 marginali $(\pi^{V=v}\pi^{Z=z}\pi^{K=k})$. La (5) può essere scritta anche nella formulazione log-lineare rappresentata nella (6)

$$\begin{cases} \log(\pi^{V=v\ Z=z\ K=k}) = u^{VZK} + \lambda^{V=v} + \lambda^{Z=z} + \lambda^{K=k} + \lambda^{V=v\ Z=z} + \lambda^{V=v\ K=k} + \lambda^{K=K\ Z=z} \\ \log(\pi^{A=a|V=v}) = u^{A|V=v} + \lambda^{A=a} + \lambda^{A=a\ V=v} \\ \log(\pi^{B=b|V=v}) = u^{B|V=v} + \lambda^{B=b} + \lambda^{B=b\ V=v} \\ \log(\pi^{C=c|Z=z}) = u^{C|\ Z=z} + \lambda^{C=c} + \lambda^{C=c\ Z=z} \\ \log(\pi^{D=d|Z=z}) = u^{D|\ Z=z} + \lambda^{D=d} + \lambda^{D=d\ Z=z} \\ \log(\ \pi^{E=e|\ K=k}) = u^{E|\ K=k} + \lambda^{E=e} + \lambda^{E=e\ K=k} \\ \log(\ \pi^{F=f|\ K=k}) = u^{F|K=k} + \lambda^{F=f} + \lambda^{F=f\ K=k} \end{cases} \qquad (6)$$

I parametri a 2 variabili delle probabilità condizionate (per esempio $\lambda^{A=a\ V=v}$) sono l'equivalente dei loadings (cioè del parametro λ_{11} della (4)), mentre quelli della congiunta (per esempio $\lambda^{V=v\ Z=z}$) sono un equivalente della covarianza.

Dalle precedenti formule (dalla (1) alla (6)) si deduce che il passaggio da una metodologia all'altra non consiste solo in un cambiamento di tipologia delle variabili, cioè da continue a categoriali, ma anche in una diversa formulazione. Nei modelli per dati continui i parametri misurano l'influenza direttamente sulla variabile risposta (infatti nella (1) e nella (4) c'è una relazione lineare tra le osservate e le latenti), nel categoriale i parametri misurano l'influenza sulla probabilità di avere un certo valore della variabile risposta (infatti nella (3) e nella (6) la variabile risposta è il logaritmo della probabilità).
Brevemente, quindi, si può facilmente implementare il CFA appena analizzato teoricamente nei 2 programmi tipici delle due metodologie: il Lisrel per i modelli SEM, il Lem per i modelli causali a classi latenti.

Input mod Lisrel: CFA

LX=FU,FI PH=SY,FR TD=DI,FR
FR LX 1 1 LX 2 1 LX 3 2 LX 4 2 LX 5 3 LX 6 3

Input mod Lem: CFA

VZK {VZ,VK,ZK}
A|V
B|V
C|Z
D|Z
E|K
F|K

dove{VZ,VK,ZK} significa che V è correlato con Z e K e K è correlato con Z

3.3.3 Errori di misura correlati

Un caso particolare di un modello CFA si presenta quando gli indicatori sono correlati sia per mezzo delle variabili latenti sia per mezzo degli errori δ. Gli errori correlati sono tipici nei modelli a dati longitudinali. Un esempio può essere l'osservare un fenomeno 2 volte, una osservazione al tempo t e una al tempo t+1. E' molto probabile che queste 2 osservazioni siano correlate fra loro per fattori diversi dalle variabili latenti che essi indicano e questa loro ulteriore relazione viene inserita nel modello attraverso la correlazione tra gli errori. Il dataset che si utilizza per paragonare i 2 metodi (SEM e modello loglineare causale a variabili latenti) è sempre quello composto da 6 variabili osservate a media nulla (A,B,C,D,E ed F) che sono raggruppabili in 3 latenti (V,Z,K) anch'esse a media nulla. Mentre nel primo paragrafo del capitolo si è ipotizzato che le 6 variabili osservate siano correlate solo attraverso le latenti, ora si ipotizza che A e C abbiano gli errori correlati.

SEM: CFA con errori correlati

$$\begin{matrix} A \\ B \\ C \\ D \\ E \\ F \end{matrix} = \begin{bmatrix} \lambda_{11} & 0 & 0 \\ \lambda_{21} & 0 & 0 \\ 0 & \lambda_{32} & 0 \\ 0 & \lambda_{42} & 0 \\ 0 & 0 & \lambda_{53} \\ 0 & 0 & \lambda_{63} \end{bmatrix} \begin{bmatrix} V \\ Z \\ K \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_5 \\ \delta_6 \end{bmatrix} \quad (7)$$

La formula (7) mostra che le variabili osservate sono a 2 a 2 indicatrici delle 3 variabili latenti, come vuole il modello teorico che si vuole testare nel CFA. I parametri λ sono i loadings. I δ sono termini di errore a media 0 e incorrelati fra loro eccetto δ_1 e δ_3 (la loro matrice di varianza –covarianza, TD, non è diagonale). Le variabili latenti (V,Z,K) sono a media nulla e correlate (la loro matrice di varianza-covarianza,PHI, è completa e simmetrica).

$$PHI = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{12} & \phi_{22} & \phi_{23} \\ \phi_{13} & \phi_{23} & \phi_{33} \end{bmatrix}$$

$$TD = \begin{bmatrix} \theta_{11} & 0 & \theta_{13} & \dots & 0 \\ 0 & \theta_{22} & 0 & \dots & \vdots \\ \theta_{13} & 0 & \theta_{33} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \dots & \theta_{66} \end{bmatrix}$$

LCA: CFA con errori correlati

Hagenaars (1988) risolve l'introduzione degli errori correlati in 2 modi

a) Ipotizzando una relazione diretta (cioè una relazione causale) fra A e C (metodo che non ha lo stesso significato in SEM):

correlazione fra gli errori: A↔C

relazione diretta tra A e C: A→C

$\pi^{A=a\ B=b\ C=c\ D=d\ E=e\ F=f\ V=v\ Z=z\ K=k} =$
$\pi^{V=v\ Z=z\ K=K}\pi^{A=a|V=v}\ \pi^{B=b|V=v}\ \pi^{C=c|Z=z\ A=a}\ \pi^{D=d|Z=z}\quad \pi^{E=e|\ K=k}\pi^{F=f|\ K=k}$
(8)

La probabilità congiunta $\pi^{A=a\ B=b\ C=c\ D=d\ E=e\ F=f\ V=v\ Z=z\ K=k}$ è scomposta nella probabilità congiunta delle latenti ($\pi^{V=v\ Z=z\ K=k}$) e nelle condizionate delle osservate ($\pi^{A=a|V=v}, \pi^{B=b|V=v}, \pi^{C=c|Z=z\ A=a}, \pi^{D=d|Z=z}, \pi^{E=e|K=k}, \pi^{F=f|\ K=k}$). La congiunta $\pi^{V=v\ Z=z\ K=k}$, non essendoci l'indipendenza tra V, Z e K , non è uguale alle 3 marginali ($\pi^{V=v}\pi^{Z=z}\pi^{K=k}$). La (8) può essere scritta anche nella formulazione log-lineare rappresentata nella (9)

$$\begin{cases} \log(\pi^{V=v\ Z=z\ K=k}) = u^{VZK} + \lambda^{V=v} + \lambda^{Z=z} + \lambda^{K=k} + \lambda^{V=v\ Z=z} + \lambda^{V=v\ K=k} + \lambda^{K=K\ Z=z} \\ \log(\pi^{A=a|V=v}) = u^{A|V=v} + \lambda^{A=a} + \lambda^{A=a\ V=v} \\ \log(\pi^{B=b|V=v}) = u^{B|V=v} + \lambda^{B=b} + \lambda^{B=b\ V=v} \\ \log(\pi^{C=c|Z=z\ A=a}) = u^{C|\ Z=z} + \lambda^{C=c} + \lambda^{C=c\ Z=z} + \lambda^{C=c\ A=a} \\ \log(\pi^{D=d|Z=z}) = u^{D|\ Z=z} + \lambda^{D=d} + \lambda^{D=d\ Z=z} \\ \log(\ \pi^{E=e|\ K=k}) = u^{E|\ K=k} + \lambda^{E=e} + \lambda^{E=e\ K=k} \\ \log(\ \pi^{F=f|\ K=k}) = u^{F|K=k} + \lambda^{F=f} + \lambda^{F=f\ K=k} \end{cases} \qquad (9)$$

I parametri a 2 variabili delle probabilità condizionate (per esempio $\lambda^{A=a\ V=v}$) sono l'equivalente dei loadings (cioè del parametro λ_{11} della (7)), mentre quelli della congiunta (per esempio $\lambda^{V=v\ Z=z}$) sono una misura della covarianza.

b) Ipotizzando la correlazione fra A e C (metodo che ha lo stesso significato in SEM):

$\pi^{A=a\ B=b\ C=c\ D=d\ E=e\ F=f\ V=v\ Z=z\ K=k} =$
$\pi^{V=v\ Z=z\ K=K}\pi^{B=b|V=v}\ \pi^{A=a\ C=c|Z=z\ V=v}\ \pi^{D=d|Z=z}\quad \pi^{E=e|\ K=k}\pi^{F=f|\ K=k}$
(10)

La probabilità congiunta $\pi^{A=a\ B=b\ C=c\ D=d\ E=e\ F=f\ V=v\ Z=z\ K=k}$ è scomposta nella probabilità congiunta delle latenti ($\pi^{V=v\ Z=z\ K=K}$) e nelle condizionate delle osservate ($\pi^{A=a|V=v}, \pi^{B=b|V=v}, \pi^{C=c|Z=z\ A=a}, \pi^{D=d|Z=z}, \pi^{E=e|K=k}, \pi^{F=f|\ K=k}$). La congiunta $\pi^{V=v\ Z=z\ K=K}$, non essendoci l'indipendenza tra V, Z e K , non è uguale alle 3 marginali ($\pi^{V=v}\pi^{Z=z}\pi^{K=k}$). La (10) può essere scritta anche nella formulazione log-lineare rappresentata nella (11)

$$
\begin{cases}
\log(\pi^{V=v\,Z=z\,K=K}) = u^{VZK} + \lambda^{V=v} + \lambda^{Z=z} + \lambda^{K=k} + \lambda^{V=v\,Z=z} + \lambda^{V=v\,K=k} + \lambda^{K=K\,Z=z} \\
\log\left(\pi^{B=b|V=v}\right) = u^{B|V=v} + \lambda^{B=b} + \lambda^{B=b\,V=v} \\
\log\left(\pi^{A=a\,C=c|Z=z\,V=v}\right) = u^{CA|\,Z=z\,V=v} + \lambda^{C=c} + \lambda^{C=c\,Z=z} + \lambda^{A=a} + \lambda^{A=a\,V=v} + \lambda^{C=c\,A=a} \\
\log\left(\pi^{D=d|Z=z}\right) = u^{D|\,Z=z} + \lambda^{D=d} + \lambda^{D=d\,Z=z} \\
\log\left(\ \pi^{E=e|\,K=k}\right) = u^{E|\,K=k} + \lambda^{E=e} + \lambda^{E=e\,K=k} \\
\log\left(\ \pi^{F=f|\,K=k}\right) = u^{F|K=k} + \lambda^{F=f} + \lambda^{F=f\,K=k}
\end{cases}
\qquad (11)
$$

I parametri a 2 variabili delle probabilità condizionate (per esempio $\lambda^{A=a\,V=v}$) sono l’equivalente dei loadings (cioè del parametro λ_{11} della (7)), mentre quelli della congiunta (per esempio $\lambda^{V=v\,Z=z}$) sono una misura della covarianza.

Comunque per poter modellare un CFA con errori correlati in LCA è preferibile utilizzare il metodo b).

Ora viene spiegato come si può implementare il CFA appena analizzato teoricamente nei 2 programmi tipici delle due metodologie: il Lisrel per i modelli SEM, il Lem per i modelli causali a classi latenti.

Input mod Lisrel: CFA

```
LX=FU,FI PH=SY,FR TD=SY,FI
FR LX 1 1 LX 2 1 LX 3 2 LX 4 2 LX 5 3 LX 6 3
FR TD 1 1 TD 2 2 TD 3 3 TD 4 4 TD 5 5 TD 6 6 TD 1 3
```

Input mod Lem: CFA

```
VZK  {VZ,VK,ZK}
B|V
AC|ZV  {AV,CZ,AC}
D|Z
E|K
F|K
```

3.4 Analisi strutturale

L’analisi strutturale è la procedura successiva all’analisi fattoriale ed è usata per determinare le relazioni esistenti tra le varie variabili latenti (nella terminologia SEM η e ξ) e quindi capire le relazioni oggetto di studio.

Il dataset che si utilizza per paragonare i 2 metodi (SEM e modello log-lineare causale a variabili latenti) è composto da 6 variabili osservate a media nulla (A,B,C,D,E ed F) che sono raggruppabili in 3 latenti (V,Z,K) anch’esse a media nulla. Nell’analisi strutturale ciò che interessa sono le relazioni tra V, Z e K: per esempio si potrebbe studiare se V influisce causalmente sia su Z che su K e se Z influisce causalmente su K, cioè un modello ricorsivo come quello rappresentato nella figura 1:

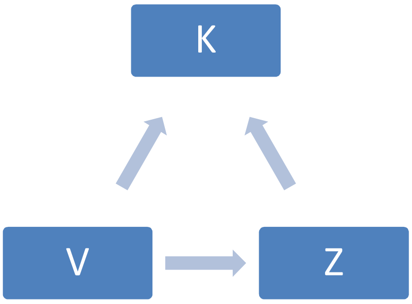


Figura 1: modello ricorsivo

L'assenza delle latenti in quest'analisi non è importante poiché quanto assunto per i modelli a variabili osservate può essere esteso a modelli con variabili latenti, naturalmente con le dovute attenzioni: l'analisi fattoriale, infatti, può portare a variabili latenti con scale differenti nelle 2 metodologie e ciò deve essere considerato quando si paragonano gli effetti strutturali. La scelta di un'analisi con i soli dati osservati è stata fatta per rendere più agevole e diretto il confronto.

SEM:

$$\begin{bmatrix} Z \\ K \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \beta_{21} & 0 \end{bmatrix} \begin{bmatrix} Z \\ K \end{bmatrix} + \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} [V] + \begin{bmatrix} \varsigma_1 \\ \varsigma_2 \end{bmatrix} \qquad (12)$$

La formula (12) mostra che la variabile V influisce direttamente su Z e su K attraverso rispettivamente i parametri γ_1 e γ_2. La variabile Z influisce su K attraverso β_{21}. I ζ sono termini di errore a media zero e incorrelati fra loro (la loro matrice di varianza – covarianza, PSI, è diagonale). Il modello che ne deriva è esattamente identificato.

$$PSI = \begin{bmatrix} \varphi_{11} & 0 \\ 0 & \varphi_{22} \end{bmatrix}$$

LCA:

$$\pi^{V=v\,Z=z\,K=k} = \pi^{V=v} \pi^{Z=z|V=v} \pi^{K=k|Z=z\,V=v} \qquad (13)$$

La probabilità congiunta $\pi^{V=v\,Z=z\,K=k}$ è scomposta nella probabilità marginale della latente ($\pi^{V=v}$) e nelle condizionate delle latenti Z e K ($\pi^{Z=z|V=v}, \pi^{K=k|Z=z\,V=v}$). La (13) può essere scritta anche nella formulazione log-lineare rappresentata nella (14)

$$\begin{cases} \log(\pi^{V=v}) = u^V + \lambda^{V=v} \\ \log(\pi^{Z=z|V=v}) = u^{Z|V=v} + \lambda^{Z=z} + \lambda^{Z=z\,V=v} \\ \log(\pi^{K=k|Z=z\,V=v}) = u^{K|V=v\,Z=z} + \lambda^{K=k} + \lambda^{K=k\,V=v} + \lambda^{K=k\,Z=z} \end{cases} \qquad (14)$$

I parametri a 2 variabili delle probabilità condizionate (per esempio $\lambda^{Z=z\,V=v}$) sono l'equivalente dei parametri β e γ della (12) (cioè del parametro γ_1 della (12)).

Utilizzando nuovamente il Lisrel per i modelli SEM, il Lem per i modelli causali a classi latenti, si ottengono i seguenti comandi:

Input mod Lisrel: CFA
BE=FU,FI PS=DI,FR GA=FU,FR
FR BE 2 1
Input mod Lem: CFA
V
Z|V
K|VZ {KV,KZ}

Una differenza tra i modelli SEM e quelli causali a classi latenti è l'impossibilità in quest'ultimi di ipotizzare modelli simultanei, cioè un modello come quello rappresentato nella figura seguente

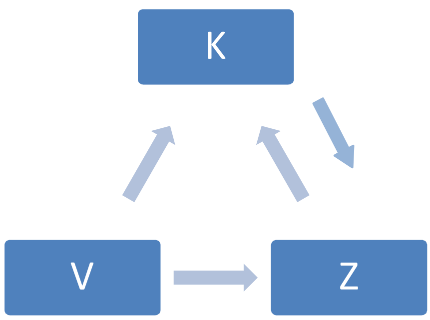


Figura 2: modello simultaneo

dove Z influisce direttamente su K che a sua volta influisce direttamente su Z. Un modello come questo in SEM è così formulato

SEM: modello simultaneo

$$\begin{bmatrix} Z \\ K \end{bmatrix} = \begin{bmatrix} 0 & \beta_{12} \\ \beta_{21} & 0 \end{bmatrix} \begin{bmatrix} Z \\ K \end{bmatrix} + \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} [V] + \begin{bmatrix} \varsigma_1 \\ \varsigma_2 \end{bmatrix} \tag{15}$$

La formula (15) mostra che la variabile V influisce direttamente su Z e su K attraverso rispettivamente i parametri γ_1 e γ_2. La variabile Z influisce su K attraverso β_{21} e la K influisce su Z attraverso β_{12}. Nel semplice caso qui considerato tale modello non è identificato, ma con una struttura più complessa è possibile stimare un modello simultaneo, ad esempio utilizzando opportune restrizioni sui parametri β e γ.

$$PSI = \begin{bmatrix} \varphi_{11} & 0 \\ 0 & \varphi_{22} \end{bmatrix}$$

LCA: dimostrazione dell'impossibilità di formulare un modello simultaneo

$$\pi^{V=v\ Z=z\ K=k} = \pi^{V=v}\pi^{Z=z\ K=k|V=v} \quad (16)$$

La probabilità congiunta $\pi^{V=v\ Z=z\ K=k}$è scomposta nella probabilità marginale della latente ($\pi^{V=v}$) e nelle condizionate delle latenti Z e K ($\pi^{Z=z\ K=k|V=v}$). Per inserire l'effetto diretto di Z su K si dovrebbe decomporre $\pi^{Z=z\ K=k|V=v}$ in $\pi^{Z=z|V=v}\pi^{K=k|V=v\ Z=z}$, per inserire l'effetto diretto di K su Z si dovrebbe decomporre $\pi^{Z=z\ K=k|V=v}$ in $\pi^{K=k|V=v}\pi^{Z=z|V=v\ K=k}$, ma queste 2 decomposizioni sono tra loro conflittuali e quindi non si riesce a formulare in maniera log-lineare un modello simultaneo.

3.5 Verifica di ipotesi sui parametri

I test sui parametri dei modelli sono possibili sia nella formulazione SEM che in quella loglineare causale per dati categoriali (o causale a classi latenti), grazie alle proprietà dei metodi di stima basati sulla ML. Se si vuole studiare se in un modello strutturale, come quello della figura 1, Z influisca su K:

SEM:

$H_0: \beta_{21} = 0\ vs\ H_1: \beta_{21} \neq 0$

grazie alle proprietà asintotiche delle stime ML, si usa un test basato sulla statistica t

t $= \frac{\beta_{21} - 0}{se\ (\beta_{21})}$

La statistica t , sotto condizioni di regolarità possedute dal modello causale in generale, si distribuisce asintoticamente come una normale con media 0 e varianza unitaria.

LCA:

$H_0: \lambda^{Z=z\ K=k} = 0\ vs\ H_1: \lambda^{Z=z\ K=k} \neq 0$ con z=1,…G e k=1,..H

che equivale a

$H_0: \tau^{Z=z\ K=k} = 1\ vs\ H_1: \tau^{Z=z\ K=k} \neq 1$ con z=1,…G e k=1,..H

Grazie alle proprietà asintotiche delle stime ML, il test che pone più vincoli contemporaneamente uguali a un valore è il test di Wald:

$H_0: \underline{\hat{\tau}} = 1\ vs\ H_1: \underline{\hat{\tau}} \neq 1$

$W_e(\underline{\tau})=(\hat{\underline{\tau}} - \underline{\tau}_0)' i(\underline{\tau}_0) (\hat{\underline{\tau}} - \underline{\tau}_0)$

Con $\underline{\tau}_0 = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ e $\hat{\underline{\tau}} = \begin{bmatrix} \tau^{Z=1\ K=1} \\ \vdots \\ \tau^{Z=G\ K=H} \end{bmatrix}$

La statistica $W_e(\underline{\tau})$, sotto alcune ipotesi di regolarità che il modello causale in generale possiede, si distribuisce asintoticamente come un χ^2 con (G-1)(H-1) gradi di libertà.

Nel caso in cui le due variabili, di cui si studia l'effetto, fossero dicotomiche (G=2, H=2), la mancanza dell'effetto potrà essere analizzata attraverso un test "normale" , cioè un test su un solo parametro (infatti in questo caso, grazie alle condizione per l'identificazione, effect code o dummy code, è solo uno il parametro da studiare), cioè

$H_0: \hat{\tau}^{Z=2\ K=2} = 1\ vs\ H_1: \hat{\tau}^{Z=2\ K=2} \neq 1$

t $=\frac{\hat{\tau}^{Z=2\ K=2} - 1}{se\ (\hat{\tau}^{Z=2\ K=2})}$

La statistica t , sotto condizioni di regolarità possedute dal modello causale in generale, si distribuisce asintoticamente come una normale con media 0 e varianza unitaria. Nel caso in cui Z fosse categoriale e K dicotomico usando il t-test, si studia non l'annullamento completo della relazione fra 2 variabili, ma solo l'indipendenza di K dalla categoria z della variabile Z, cioè non si ha più l'indipendenza ma la quasi indipendenza (Hagenaars 1990).

Il t-test viene inserito nell'output sia del LEM che del Lisrel, il test di Wald solo nell'output del LEM, perché in Lisrel il test di Wald coincide con il t-test.

3.5.1 Test sul modello completo

Esistono vari test per scegliere un modello e studiarne la buona adattabilità ai dati. I test che possono essere usati per studiare la bontà di un modello in entrambe le metodologie (SEM e log-lineare causale) sono :

1) Test chi-quadrato (test di Pearson), che verifica la conformità ai dati del modello ipotizzato.

$$X^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

dove O_i è la frequenza osservata e E_i è quella attesa. La statistica X^2, sotto l'ipotesi nulla, è distribuita come un χ^2 con gradi di libertà che dipendono dal modello stesso.

2) BIC

Il BIC tiene in considerazione sia l'adattabilità dei dati al modello sia il numero di parametri, penalizzando i modelli con più parametri

BIC=$-2\ln(L) + k\ln(n)$

dove k è il numero dei parametri del modello, L è il valore massimizzato della funzione di verosimiglianza ed n il numero di osservazioni. Si sceglie il modello con BIC minore.

3) AIC

L'AIC tiene in considerazione sia l'adattabilità dei dati al modello sia il numero di parametri, penalizzando i modelli con più parametri, ma meno del BIC

AIC=$2k - 2\ln(L)$

dove k è il numero dei parametri del modello e L è il valore massimizzato della funzione di massima verosimiglianza. Si sceglie il modello con AIC minore.

3.6 Effetti diretti, indiretti e totali

In questo paragrafo si confronta la stima fatta con il metodo SEM con quella fatta con il metodo "modello loglineare causale". Si sceglie di utilizzare un dataset senza variabili latenti solo per semplicità, ma i risultati si applicano ugualmente coniugando l'analisi causale con il modello di misura. Vi sono solo ulteriori problemi di interpretazione dovuti però solo a problemi di identificazione "banali". L'analisi fattoriale, ad esempio, fornisce spesso variabili latenti con segno opposto nelle 2 metodologie (SEM ed LCA), rendendo il confronto meno evidente. Per esempio una variabile latente che indica il reddito futuro atteso può essere ordinata decrescente in LCA (classe 1: alto reddito futuro, classe 2: basso reddito futuro) e crescente in SEM (la variabile reddito futuro atteso è crescente). La differenza nelle variabili latenti deve essere ricordata quando si confrontano i parametri strutturali. Ritornando all'esempio del reddito futuro atteso e studiando come esso influisca sul consumo corrente (ipotizzato crescente sia in SEM che in LCA), risulta che un parametro strutturale positivo in SEM è equivalente come interpretazione ad un parametro negativo in LCA. Se, invece, il consumo fosse decrescente in LCA e crescente in SEM, risulterebbe che un parametro strutturale positivo in SEM corrisponderebbe ad un parametro positivo in LCA.

I dati utilizzati sono quelli generati nel capitolo 1 : per il modello "completo" si utilizzano i dati della tabella 14, per il modello "con solo l'effetto indiretto" quelli della tabella 16, per il modello "primo caso di presenza del solo effetto diretto" quelli della tabella 18 e per il modello "secondo caso di presenza del solo effetto diretto" quelli della tabella 20. I dati per il modello log-lineare causale sono quelli presenti nelle tabelle succitate, mentre per il modello SEM richiedono un'ulteriore operazione: si devono creare dalle tabelle le matrici di varianza- covarianza. Per semplicità si ipotizzerà che le 3 variabili dicotomiche (A,B,C) possano prendere i valori 0 o 1, così che:

E(A)=1P(A=1)+0 P(A=0)= P(A=1)

E(B)=1P(B=1)+0 P(B=0)= P(B=1)

E(C)=1P(C=1)+0 P(C=0)= P(C=1)

I valori da inserire nella matrice di varianza-covarianza allora divengono

Var(A)= $(1 - P(A = 1))^2$P(A=1)+ $(0 - P(A = 1))^2$P(A=0)

Var(B)= $(1 - P(B = 1))^2$P(B=1)+ $(0 - P(B = 1))^2$P(B=0)

Var(C)= $(1 - P(C = 1))^2$P(C=1)+ $(0 - P(C = 1))^2$P(C=0)

Cov(A,B)=P(A=1,B=1)-P(A=1)P(B=1)

Cov(A,C)=P(A=1,C=1)-P(A=1)P(C=1)

Cov(C,B)=P(C=1,B=1)-P(C=1)P(B=1)

La covarianza Cov(A,B) è ottenuta, con semplici formule, dalla

E(A,B)-E(A)E(B)

Lo stesso procedimento è utilizzato per Cov(A,C) e Cov (C,B).

Come spiegato in precedenza, in tutti gli esempi seguenti non si analizza l'inferenza in quanto i dati LCA sono "esatti".

3.6.1 Modello completo

Il modello completo è quello in cui A influisce sia su B che su C e A influisce su B. In un modello così esistono 3 effetti diretti (A→B, A→C e B→C) e un effetto indiretto (A→B→C, cioè A influisce su C attraverso B). La matrice di varianza-covarianza calcolata a partire dalla tabella 14 diviene:

Tabella 5: matrice varianza-covarianza modello completo

	A	B	C
A	0.0554		
B	-0.0322	0.2158	
C	-0.0058	0.0475	0.1302

Ora si confrontano le stime dei parametri con le 2 metodologie (SEM e modelli causali loglineari):

Tabella 6:confronto effetti metodo SEM-loglineare causale

	SEM	Modello causale log-lineare
Effetto diretto di A su B	-0.58	-0.6931
Effetto diretto di A su C	0.03	0.1823
Effetto diretto di B su C	0.22	1.1631

Effetto totale di A su C	-0.10	-0.2917
Effetto indiretto totale di A su C	-0.13	-0.4740
Effetto cella		-0.3518

Risulta evidente che i segni di tutti gli effetti coincidono nelle 2 metodologie: A influisce negativamente su B e positivamente su C, B influisce positivamente su C. L'effetto totale di A su C è negativo a causa dell'effetto indiretto, la cui intensità è tale da coprire completamente la positività dell'effetto diretto

effetto indiretto: A↑ =>B↓=> C↓

effetto diretto: A↑=>C↑

effetto totale= effetto diretto + effetto indiretto; |effetto indiretto|> effetto diretto=> prevale segno effetto indiretto

3.6.2 Modello "solo effetto indiretto"

Il modello "solo effetto indiretto" è quello in cui A influisce solo su B e B influisce su C. In un modello così esistono 2 effetti diretti (A→B e B→C) e un effetto indiretto (A→B→C, cioè A influisce su C attraverso B). La matrice di varianza-covarianza diviene:

Tabella 7: matrice varianza-covarianza modello "solo effetto indiretto"

	A	B	C
A	0.0554		
B	-0.0322	0.2158	
C	-0.0092	0.0619	0.1601

Ora si confrontano le stime dei parametri con le 2 metodologie (SEM e modelli causali loglineari):

Tabella 8:confronto effetti metodo SEM-loglineare causale

	SEM	Modello causale log-lineare
Effetto diretto di A su B	-0.58	-0.6931
Effetto diretto di A su C	0	0
Effetto diretto di B su C	0.29	1.1631
Effetto totale di A su C	-0.17	-0.4419
Effetto indiretto totale di A su C	-0.17	-0.4419
Effetto cella		0

Risulta evidente che i segni di tutti gli effetti coincidono nelle 2 metodologie: A influisce negativamente su B e B influisce positivamente su C. L'effetto totale di A su C, che coincide con l'effetto indiretto totale, è negativo effetto indiretto: A↑ =>B↓=> C↓

3.6.3 Modello "primo caso solo effetto diretto"

Il modello "primo caso solo effetto indiretto" è quello in cui A influisce solo su C e B influisce su C. In un modello così esistono 2 effetti diretti (A→C e B→C) e nessun effetto indiretto, ma nella metodologia loglineare causale rimane l'effetto cella . La matrice di varianza-covarianza diviene:

Tabella 9: matrice varianza-covarianza modello "primo caso solo effetto diretto"

	A	B	C
A	0.0554		
B	0	0.238	
C	0.0033	0.0541	0.0832

Ora si confrontano le stime dei parametri con le 2 metodologie (SEM e modelli causali loglineari):

Tabella 10:confronto effetti metodo SEM-loglineare causale

	SEM	Modello causale log-lineare
Effetto diretto di A su B	0	0
Effetto diretto di A su C	0.06	0.1823
Effetto diretto di B su C	0.23	1.1631
Effetto totale di A su C	0.06	0.1471
Effetto indiretto totale di A su C	0	-0.0352
Effetto cella		-0.0352

Risulta evidente che i segni di tutti gli effetti coincidono nelle 2 metodologie: A influisce positivamente su C, B influisce positivamente su C. L'effetto totale di A su C nel modello SEM coincide con l'effetto diretto, nel modello loglineare causale con la sommatoria tra l'effetto diretto e l'effetto cella (0.1471=0.1823-0.0352).

3.6.4 Modello "secondo caso solo effetto diretto"

Il modello "secondo caso solo effetto indiretto" è quello in cui A influisce solo su C e su B. In un modello così esistono 2 effetti diretti (A→C e A→B) e nessun effetto indiretto. La matrice di varianza-covarianza diviene:

Tabella 11: matrice varianza-covarianza modello "secondo caso solo effetto diretto"

	A	B	C
A	0.0554		
B	-0.0322	0.2158	
C	0.0015	-0.0009	0.0278

Ora si confrontano le stime dei parametri con le 2 metodologie (SEM e modelli causali loglineari):

Tabella 12:confronto effetti metodo SEM-loglineare causale

	SEM	Modello causale log-lineare
Effetto diretto di A su B	-0.58	-0.6931
Effetto diretto di A su C	0.03	0.1823
Effetto diretto di B su C	0	0
Effetto totale di A su C	0.03	0.1823
Effetto indiretto totale di A su C	0	0
Effetto cella		0

Si rileva che i segni di tutti gli effetti coincidono nelle 2 metodologie: A influisce negativamente su B e positivamente su C. L'effetto totale di A su C coincide nelle 2 metodologie con l'effetto diretto.

3.6.5 Effetto totale nullo

Il modello con effetto totale nullo è quello in cui A influisce sia su B che su C e A influisce su B, ma l'effetto totale di A su C è nullo. In un modello così esistono 3 effetti diretti (A→B, A→C e B→C) e un effetto indiretto (A→B→C, cioè A influisce su C attraverso B). E' interessante analizzare con le 2 metodologie (SEM e log-lineare causale) un modello come questo per vedere se l'effetto totale, combinazione dei 3 effetti diretti, è nullo per entrambe. Si crea un nuovo dataset affinché la combinazione dei 3 effetti diretti porti ad un effetto totale nullo, quindi, rispetto al dataset della tabella 14 del capitolo 1, si modificano i parametri "effetto diretto" τ^{AB}, τ^{AC} e τ^{CB}, poiché per questi dati nel capitolo 1 non era stato costruito un modello ad effetto nullo. I dati sono generati dai seguenti parametri:

$$\tau^{A} = 0.25$$

$$\tau^{B} = 0.8$$
$$\tau^{C} = 0.2$$

$$\tau^{AB} = 0.5577$$

$$\tau^{AC} = 1.7319$$

$$\tau^{CB} = 3.9$$

Rispetto al modello completo si cambiano τ^{AB}, τ^{AC} e τ^{CB} in modo tale che l'effetto totale sia nullo. Per costruire i dati si utilizza il seguente ordine: prima si calcola P(A), poi la condizionata P(B|A) e successivamente P(C|AB). Moltiplicando P(A), P(B|A) e P(C|AB) si ottiene la congiunta P(CAB) che è il dataset utilizzato nella stima

Tabella 13:condizionata di B, di C e congiunta CAB, sub modello "effetto totale nullo"

b|a		c|ab		cab	
1|1	0.1660	1|11	0.6460	111	0.0062
2|1	0.8340	2|11	0.3540	112	0.0004
1|2	0.6730	1|12	0.0078	121	0.1068
2|2	0.3270	2|12	0.9922	122	0.0003
		1|21	0.1686	211	0.0035
		2|21	0.8314	212	0.0487
		1|22	0.0009	221	0.5266
		2|22	0.9991	222	0.3075
				Tot	1
$\eta^{B\|A=1} = 0.3721$ $\eta^{B\|A=2} = 0.4691$		$\eta^{C\|A=1\ B=1}$ =0.4782 $\eta^{C\|A=1\ B=2}$ =0.0881 $\eta^{C\|A=2\ B=1}$ =0.3744 $\eta^{C\|A=2\ B=2}$ =0.0296			

La matrice di varianza-covarianza diviene:

Tabella 14: matrice varianza-covarianza modello completo

	A	B	C
A	0.0554		
B	-0.0281	0.2295	
C	0	0.0399	0.1008

Ora si confrontano le stime dei parametri con le 2 metodologie (SEM e modelli causali loglineari):

Tabella 15:confronto effetti metodo SEM-loglineare causale

	SEM	Modello causale log-lineare
Effetto diretto di A su B	-0.51	-0.5839
Effetto diretto di A su C	0.09	0.5492
Effetto diretto di B su C	0.19	1.3610
Effetto totale di A su C	0	0
Effetto indiretto totale di A su C	-0.09	-0.5492
Effetto cella		-0.1537

Si osserva che i segni di tutti gli effetti coincidono nelle 2 metodologie: A influisce negativamente su B e positivamente su C, B influisce positivamente su C. L'effetto totale di A su C è nullo a causa dell'effetto indiretto la cui intensità è tale da coprire completamente la positività dell'effetto diretto

effetto indiretto: A↑ =>B↓=> C↓

effetto diretto: A↑=>C↑

effetto totale= effetto diretto + effetto indiretto; |effetto indiretto|= effetto diretto=> effetto totale nullo

3.6.6Conclusioni

Nei modelli appena studiati, utilizzare la metodologia SEM o quella loglineare causale porta a risultati sostanzialmente equivalenti per quanto riguarda la presenza/assenza di una relazione positiva/negativa tra le variabili. Questo è un punto di partenza importante perché, essendo dati "esatti", si conosce il vero modello e quindi si può affermare che entrambe le metodologie identificano i parametri in modo giusto. Si mette in evidenza soprattutto che la metodologia SEM riesce a riconoscere i sub modelli (solo effetto indiretto, effetto totale nullo, etc) senza imporre vincoli anche se i dati sono creati con la logica loglineare.

3.7 Conclusioni confronto fra modello SEM e modello causale a variabili latenti

In questo capitolo sono stati messi a confronto il modello SEM e quello loglineare causale a classi latenti. La tabella 16, che sintetizza il capitolo, indica che in via generale i due approcci hanno le stesse caratteristiche, ma con problematiche differenti:

Tabella 16: comparazione modello causale loglineare a classi latenti-sem

	Mod. causale loglinere a classi latenti	Sem
ANALISI FATTORIALE		
NUMERO FATTORI	X	X
CORRELAZIONE FRA ERRORI	X	X
NUMERO CLASSI	X	
ANALISI STRUTTURALE		
EFFETTI DIRETTI	X	X
Significatività parametri	X	X
EFFETTI TOTALI	X	X
Significatività parametri		X
EFFETTI INDIRETTI	X	X
Significatività parametri		X

CORRELAZIONE ETA	X	X
Significatività parametri	X	X
MODELLO TOTALE		
TEST (AIC, CHI-QUADRO…)	X	X
MODELLO AD EQUAZIONI SIMULTANEE		X

Con il modello causale si hanno molti più problemi a stimare modelli complessi che con il SEM, a causa delle celle sparse. All’aumentare delle variabili, il numero delle celle aumenta moltissimo, per esempio con 2 variabili dicotomiche, le celle divengono 4, con 3 variabili, le celle divengono 8. Un'altra causa della presenza delle celle sparse può essere un elevato numero di categorie delle variabili. Questo aumento delle celle in campioni di piccole dimensioni può portare alla formazioni di celle vuote con la conseguenza che la distribuzione chi_quadro non può essere usata per calcolare il p-value, perché L^2 non è ben approssimato da tale distribuzione. Per risolvere questo problema, si stima il p-value con l’approccio bootstrap. Un altro problema, che si riscontra, è quello del locale massimo che si presenta quando la massimizzazione della funzione di verosimiglianza non è l’unica, ma ce ne sono molte. Un ulteriore problema del modello LCA nasce dai valori” boundary”: le probabilità stimate 0 (valori “boundary”) influiscono sul numero di gradi di libertà del L^2.

Capitolo 4:
esempi confronto Lisrel-Lem

4.1 Un'applicazione empirica

Gli esempi seguenti mostrano come anche in un'analisi empirica con dati reali l'analisi SEM e l'analisi loglineare causale a classi latenti portino agli stessi risultati sostanziali. I dati sono presi da una survey proposta ai partecipanti alla festa del prosciutto di Sauris. L'applicazione dei modelli SEM a questo dataset è già stata analizzata da Pistore (2010) al quale si rinvia per ulteriori dettagli. Durante la festa, per le vie del borgo alpino, si sviluppa un'intensa attività con il mercatino dei prodotti agro-alimentari carnici e dell'artigianato locale (lavorazione tessile e del legno); comunque "Il re incontrastato della festa" rimane il prosciutto, che può essere degustato ed acquistato. Certificato IGP, è prodotto solo da 2 aziende site nel comune omonimo, secondo un procedimento che rende unica la leggera affumicatura: ottenuta bruciando esclusivamente legno di faggio, di cui sono ricchi i boschi locali, in appositi camini la cui canna fumaria convoglia il fumo nelle stanze adibite a tale processo.
Lo studio, che può essere fatto sui dati di questa survey, rientra nell'ambito della "soddisfazione del cliente". Nel campo della ricerca di marketing la soddisfazione del cliente, che misura quanto i prodotti e/o i servizi offerti ai clienti incontrino le loro aspettative, gioca un ruolo importante nella scelta delle politiche da attuare per migliorare la propria competitività nel mercato. Riuscire ad aumentare la soddisfazione del cliente vuol dire guadagnare fette di mercato a danno dei concorrenti e quindi maggiori introiti. Anderson e Mitall (2000) sostengono che la "soddisfazione" crea un guadagno economico particolarmente quando la performance del servizio offerto dall'impresa eccede l'aspettativa del cliente. In economia del turismo (Cronin et al. 2000, Chen & Tsai 2007), si sostiene che la soddisfazione è quindi una delle variabili che determinano la behavioral intention futura. Come osservato da Chen & Tsai (2007) comprendere le relazioni tra i comportamenti futuri e i suoi " fattori determinanti" può migliorare lo sforzo fatto per massimizzare il guadagno sulle risorse a disposizione. Si ipotizza, come è stato dimostrato in molti studi (Chen & Tsai 2009, Cronin et al 2000, Baker& Crompton 2000) , che le variabili esogene del modello siano la qualità, l'interesse e le caratteristiche sia del prodotto che della festa e che le variabili endogene

siano la soddisfazione e il comportamento futuro. In un modello strutturale così, per esempio, è possibile verificare se l'interesse verso la manifestazione influisca sul comportamento futuro nei riguardi del prosciutto e della festa solo attraverso la soddisfazione (effetto indiretto) o anche direttamente (effetto diretto), così da capire se un'operazione di marketing che aumenti l'interesse del visitatore possa aumentare la vendita del prodotto e l'affluenza alla manifestazione.

L'oggetto di studio è come la festa di Sauris influisca sui comportamenti futuri (behavioral intention) nei riguardi della medesima festa e del prosciutto omonimo. Per comportamento futuro si intende, in generale in economia del turismo, l'idea del visitatore di rivisitare una destinazione o la volontà di raccomandare la destinazione ad altri: in questo specifico caso, l'idea di ritornare alla festa di Sauris o di consigliarla ad altri. Il comportamento futuro è applicato anche al prodotto e quindi rappresenta la volontà di riacquistare o di consigliare il prosciutto di Sauris.
Per soddisfazione si intende il piacere e la contentezza provati dal visitatore durante la festa del prosciutto di Sauris. La soddisfazione può essere vista come la capacità della manifestazione di assecondare i desideri, le aspettative e i bisogni dei partecipanti alla festa.
Il questionario è composto da domande con scala a 7 risposte multiple ed è questo il numero delle categorie delle variabili per il modello SEM. Per il causale, si è preferito dicotomizzare le variabili in base alla media aritmetica: viene dato valore 1 alle categorie che stanno sotto la media e 2 a quelle che stanno sopra la media. Le domande sono raggruppate in 14 gruppi. Fra le variabili di potenziale interesse il gruppo "7" si riferisce all'interesse verso la manifestazione, il gruppo "8" ai giudizi sulla manifestazione, il "9" e il "10" alle determinanti l'acquisto del prosciutto crudo, l'"11" all'opinione sul prosciutto crudo, il "12" alla soddisfazione sulla festa, il "13" ai comportamenti futuri nei riguardi del prosciutto e della festa di Sauris.
Nel seguito si confrontano i risultati di SEM e LCA su semplici modelli con poche variabili estratte dal questionario. La generalizzazione a modelli più complessi porta a risultati equivalenti ma è resa difficile dalla numerosità ridotta di interviste (348).
Prima si analizza la CFA e poi la parte strutturale, la cui bontà verrà analizzata in base al BIC e all'AIC, per ovviare al problema delle celle sparse. La significatività dei parametri viene analizzata sia nel SEM che nel causale con un t-test. Il valore dei parametri nei 2 metodi non è confrontabile, perché si riferiscono a "misure" diverse: nel SEM i parametri sono l'influenza diretta sulla variabile risposta, nel loglineare causale a classi latenti sulla probabilità della variabile risposta. Per questo motivo si confronterà solo il segno del parametro, cioè se c'è un effetto positivo o negativo, e la sua presenza, cioè se il parametro è uguale 0.

4.2 Caso 1:Analisi fattoriali esplorative

Per confrontare un'analisi esplorativa nelle 2 metodologie (SEM e LCA) si scelgono 8 quesiti dal gruppo "12" del questionario (soddisfazione della festa). Si utilizzano come variabili osservate le impressioni , misurate su una scala da 1 a 7, sulle affermazioni seguenti:

1)pensare alla festa del prosciutto di Sauris mi rende felice (SA)
2)la festa del prosciutto di Sauris mi dà un senso di gioia (SE)
3)la qualità della festa del prosciutto di Sauris è notevole (SM)
4)provo una piacevole sensazione quando penso alla festa del prosciutto di Sauris (SN)
5)la festa del prosciutto di Sauris è una manifestazione che gode di una buona reputazione (SO)
6)sono soddisfatto di essere venuto alla festa del prosciutto di Sauris (SP)
7)la festa del prosciutto di Sauris è un evento ben organizzato (SQ)
8)la qualità dei prodotti della festa del prosciutto di Sauris è molto affidabile (SR)

Le domande 1-2-4-5-6 misurano per definizione la soddisfazione (Oliver 1997). La 3, la 7 e la 8 rientrano come misura della soddisfazione e non della qualità se si prende in considerazione il paper di Baker & Crompton (2000), in cui viene ben analizzata la linea di confine tra ciò che misura la soddisfazione e ciò che misura la qualità . I 2 autori, infatti, definiscono le qualità della performance come le caratteristiche del servizio che sono controllate dall'esercente e la soddisfazione come lo stato emozionale della mente influenzato da condizioni sociali e psicologiche del turista, derivate dal luogo (umore, disposizione, bisogni) e da fattori esterni che sono fuori dal controllo dell'esercente (clima, interazioni sociali, etc).

Nell'analisi SEM si usa la variabile osservata a 7 classi ordinate, in Lem una dicotomizzazione della variabile a 7 classi (classe 1: valori che stanno sotto la media, classe 2 valori che stanno sopra la media)

SEM

Come primo passaggio si calcola la matrice campionaria di varianza-covarianza; poi si inizia l'analisi ipotizzando di avere un modello ad un fattore e lo si stima col metodo MINRES (i loadings sono calcolati con il programma Lisrel, i test con R):

Tabella 12: loadings modello ad 2 fattori

Unrotated Factor Loadings

	Factor 1	Factor 2	Unique Var
SA	0.823	0.000	0.322
SE	0.856	0.095	0.259
SM	0.570	0.563	0.358
SN	0.811	0.181	0.310
SO	0.573	0.506	0.416
SP	0.589	0.548	0.353
SQ	0.459	0.728	0.260
SR	0.496	0.649	0.334

Il test (cioè quello che compara la statistica del modello saturo con quella del modello a 2 fattori) ha un p-value di 1(U=0.07, $d_f = \frac{(8-2)^2-10}{2}$=13) , che è un buon modello .

Tabella 12: loadings modello ad 1 fattore

Unrotated Factor Loadings

	Factor 1	Unique Var
	--------	----------
SA	0.663	0.560
SE	0.747	0.441
SM	0.790	0.377
SN	0.770	0.406
SO	0.763	0.418
SP	0.799	0.362
SQ	0.762	0.420
SR	0.762	0.419

Il test che compara la statistica del modello ad un fattore con quella del modello saturo ha un p-value di 1 (U=0.61, $d_f = \frac{(8-1)^2-9}{2}$=20), perciò è buon modello. Con l'analisi della statistica U si sceglie il modello ad 1 fattore perché Bartholomew (1997) consiglia di procedere a diminuire il numero di fattori finché il modello non devvii significativamente dai dati. Il modello ad 1 fattore non devia dai dati.

LCA

Come primo passaggio si calcola un modello ad 1 fattore. Il BIC risultate è 4141, il test di bontà di adattamento (36.6 df 240) ha un p-value uguale a 1. Si può già accettare un modello ad 1 fattore. Si calcola un modello a 2 fattori. Il BIC risultate è 4190, il test di bontà di adattamento (32.6 df233) ha un p-value uguale a 1 e quindi è un buon modello. La scelta del numero di fattori avviene o scegliendo il primo modello "buono" o quello con minor BIC. Con entrambi i metodi si sceglie il modello ad 1 fattore.

4.3 Caso 2: modello ricorsivo completo

In questo modello si cerca di capire come l'interesse verso il prosciutto di Sauris possa influire sui comportamenti futuri riguardanti il prodotto. Si vede, dall'analisi del modello, che all'aumentare dell'interesse, aumenta il desiderio di riacquistare il prosciutto e di consigliarlo ad altri. L'interesse influisce in 2 modi sul comportamento futuro: sia direttamente che indirettamente, cioè attraverso la soddisfazione emozionale. I partecipanti più interessati al prosciutto risultano essere anche i più soddisfatti dalla festa e conseguentemente anche più propensi a riacquistare e a raccomandarne l'acquisto ad altri. La festa, quindi, aumenta la propensione all'acquisto del prodotto e la sua pubblicizzazione (intesa come passaparola). Le variabili osservate di questo sottomodello sono MI e ML, indicatori dell'interesse verso il prosciutto, SA e SE, indicatori della soddisfazione e BF e BG, indicatori del comportamento futuro verso il prosciutto.

4.3.1 CFA

Per poter confrontare la metodologia LCA e quella SEM si scrive lo stesso modello con le 2 metodologie:

SEM

$$\begin{bmatrix} MI \\ ML \\ SA \\ SE \\ BF \\ BG \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ \lambda_2 & 0 & 0 \\ 0 & \lambda_3 & 0 \\ 0 & \lambda_4 & 0 \\ 0 & 0 & \lambda_5 \\ 0 & 0 & \lambda_6 \end{bmatrix} \begin{bmatrix} X \\ K \\ V \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix}$$

$$\text{TD}=\begin{bmatrix} \theta_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & \theta_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & \theta_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \theta_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & \theta_{66} \end{bmatrix} \quad \text{PH}=\begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{12} & \phi_{22} & \phi_{23} \\ \phi_{13} & \phi_{23} & \phi_{33} \end{bmatrix}$$

LCA

$$\pi^{X=x\ K=k\ V=v\ MI=mi\ ML=ml\ SA=sa\ SE=se\ BF=bf\ BG=bg}$$
$$= \pi^{X=x\ K=k\ V=v} \pi^{MI=mi|X=x} \pi^{ML=ml|X=x} \pi^{SA=sa|K=k} \pi^{SE=se|K=k} \pi^{BF=bf|V=v} \pi^{BG=bg|V=v}$$

Come metodo di identificazione nel modello LCA si utilizza l'effect code e per il metodo SEM si sceglie di porre la varianza della X uguale ad 1 e i loadings di K su SA e di V su BF uguali a 1. Sono stati scelti 2 metodi diversi di identificazione tra la variabile X e le variabili K e V in quanto nella parte strutturale del modello X è esogena e K e V sono endogene.

- MATRICE LX

<table>
<tr>
<td rowspan="2">L
C
A</td>
<td>X</td>
<td>1=
MOLT
INT
2=
POCO
INT</td>
<td>* TABLE X.MI [or P(MI|X)] *
Effect lambda std err z-value exp(lambda) Wald df prob
MI
1 -0.7008 0.4170 -1.681 0.4962 2.82 1 0.093
X.MI
1 1 -1.0409 0.3911 -2.662 0.3531 7.08 1 0.008
* TABLE X.ML [or P(ML|X)] *
Effect lambda std err z-value exp(lambda) Wald df prob
ML
1 -0.3737 0.1867 -2.002 0.6882 4.01 1 0.045
X.ML
1 1 -0.8024 0.1645 -4.877 0.4482 23.78 1 0.000

MI e ML sono variabili dicotomiche che rappresentano rispettivamente “l’interesse nell’approfondire la conoscenza del prosciutto di Sauris” e “l’interesse alla possibilità di assaggiare il prodotto”. La classe 1 significa poco interesse, la classe 2 molto interesse; poiché alla classe X=1 è associata in entrambi i casi la classe 2 e viceversa, ciò significa che la variabile latente X ha come valore 1 il molto interesse verso il prodotto e come valore 2 il poco interesse.</td>
</tr>
<tr>
<td>K</td>
<td>1=
MOLT
SOD
2=
POCA
SOD</td>
<td>* TABLE K.SA [or P(SA|K)] *
Effect lambda std err z-value exp(lambda) Wald df prob
SA
1 -0.0513 0.1642 -0.313 0.9500 0.10 1 0.755
K.SA
1 1 -0.9913 0.1377 -7.201 0.3711 51.85 1 0.000
* TABLE K.SE [or P(SE|K)] *
Effect lambda std err z-value exp(lambda) Wald df prob
SE
1 0.0292 0.1469 0.199 1.0296 0.04 1 0.842
K.SE
1 1 -0.9281 0.1242 -7.475 0.3953 55.88 1 0.000

SA e SE sono variabili dicotomiche che rappresentano rispettivamente “pensare alla festa del prosciutto di Sauris mi rende felice” e “la festa del prosciutto di Sauris mi dà un senso di gioia”. La classe 1 significa poca felicità/gioia, la classe 2 molta; poiché alla classe K=1 è associata in entrambi i casi la classe 2 e viceversa, ciò significa che la variabile latente K ha come valore 1 la molta soddisfazione sulla festa, come valore 2 la poca.</td>
</tr>
</table>

<table>
<tr><td></td><td>V</td><td>1=
COMP
MOLT
POS
2=
COMP
POCO
POS</td><td>* TABLE V.BF [or P(BF|V)] *
effect lambda std err z-value exp(lambda) Wald df prob
BF
1 0.0067 0.1530 0.044 1.0067 0.00 1 0.965
V.BF
1 1 -0.9174 0.1375 -6.673 0.3995 44.53 1 0.000
* TABLE V.BG [or P(BG|V)] *
effect lambda std err z-value exp(lambda) Wald df prob
BG
1 -0.0479 0.0963 -0.498 0.9532 0.25 1 0.619
V.BG
1 1 -0.6397 0.0895 -7.149 0.5275 51.11 1 0.000

BF e BG sono variabili dicotomiche che rappresentano rispettivamente “parlerò in modo positivo del prosciutto di Sauris” e “acquisterò il prosciutto di Sauris più spesso”. La classe 1 significa che non si parlerà/non si comprerà , la classe 2 che si comprerà/si parlerà; poiché alla classe V=1 è associata in entrambi i casi la classe 2 e viceversa, ciò significa che la variabile latente V ha come valore 1 un comportamento futuro positivo nei confronti del prosciutto di Sauris, come valore 2 un comportamento negativo o meno positivo.</td></tr>
</table>

<table>
<tr><td>L
I
S
R
E
L</td><td>X</td><td>↑ INT</td><td>LAMBDA-X

K V X
-------- -------- --------
MI - - - - 2.77
(0.19)
14.93
ML - - - - 2.93
(0.21)
14.16</td></tr>
<tr><td></td><td>K</td><td>↑
SODD</td><td>LAMBDA-X

K V X
-------- -------- --------
SA 1.00 - - - -
SE 0.54 - - - -
(0.04)
13.87</td></tr>
<tr><td></td><td>V</td><td>↑
COMP</td><td>LAMBDA-X

K V X
-------- -------- --------
BF - - 1.00 - -
BG - - 3.37 - -
(0.35)
9.62</td></tr>
</table>

Nel modello SEM X, K e V sono variabili crescenti, in LCA X, K e V sono decrescenti.

- MATRICE PHI

Mentre nel caso SEM i parametri sono effettivamente le Varianze-Covarianze delle variabili, nel modello LCA a classi latenti questo non è così: essi sono i parametri che definiscono la distribuzione congiunta e misurano effettivamente la relazione tra le variabili, ma non corrispondono numericamente alle varianze-covarianze (infatti se il parametro moltiplicativo è 1 o il log-lineare è 0, l’effetto non c’è e le variabili sono indipendenti come nel caso di covarianza nulla). La distribuzione congiunta è infatti:

$$\pi^{XVK} = \eta\tau^{X}\tau^{V}\tau^{K}\tau^{XK}\tau^{XV}\tau^{VK}$$

LCA	X	K	X.K effect lambda std err z-value exp(lambda) Wald df prob 1 1 0.4779 0.1501 3.183 1.6126 10.13 1 0.001
		V	X.V effect lambda std err z-value exp(lambda) Wald df prob 1 1 0.5105 0.1858 2.748 1.6662 7.55 1 0.006
	K	V	K.V effect lambda std err z-value exp(lambda) Wald df prob 1 1 0.4999 0.1408 3.551 1.6485 12.61 1 0.000

LISREL	X	X	1
		K	1.88 (0.21) 8.84
		V	1.01 (0.11) 8.76
	K	K	10.53 (1.26) 8.35
		V	3.88 (0.47) 8.26
	V	V	3.42 (0.43) 7.99

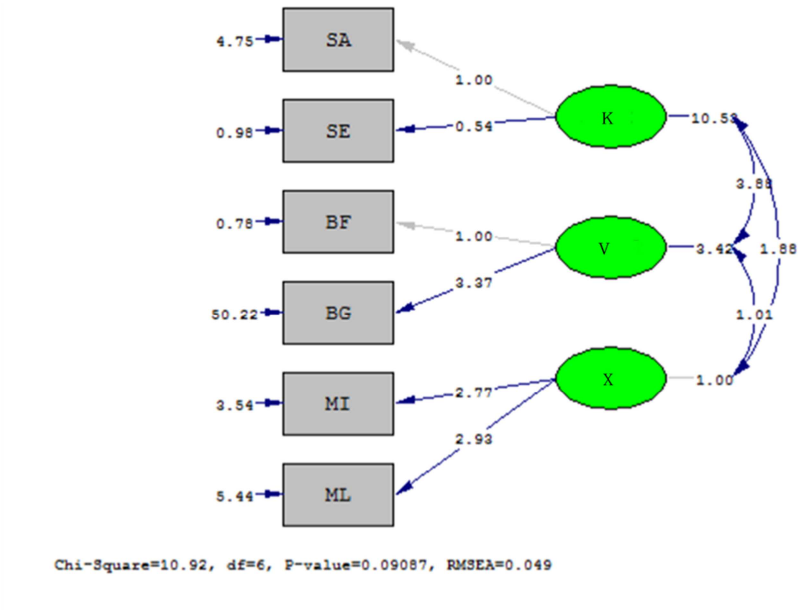


Figura 1: path diagram CFA modello interesse-soddisfazione-comportamento futuro su prosciutto

4.3.2 Analisi strutturale

Si parte da un modello completo, nel quale tutti i γ e tutti i β, in termini di Lisrel, sono compresi. Come per il modello SEM, anche per il modello loglineare causale a classi latenti gli indici di bontà del modello CFA e di quello completo rimangono gli stessi, questo perché porre una correlazione o un effetto diretto non cambia la bontà del modello. Per poter confrontare la metodologia LCA e quella SEM si scrive lo stesso modello con le 2 metodologie:

SEM

$$\begin{bmatrix} K \\ V \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} \mathrm{X} + \begin{bmatrix} 0 & 0 \\ \beta_{2\,1} & 0 \end{bmatrix} \begin{bmatrix} K \\ V \end{bmatrix} + \begin{bmatrix} \varsigma_1 \\ \varsigma_2 \end{bmatrix} \quad \mathrm{PSI}=\begin{bmatrix} \psi_{11} & 0 \\ 0 & \psi_{22} \end{bmatrix} \quad \mathrm{PH}=[\phi_{11}]$$

LCA

$$\pi^{X=x\;K=k\;V=v} = \pi^{X=x}\pi^{K=k|X=x}\pi^{V=v|X=x\;K=k}$$

In questo caso il modello completo è il miglior modello in entrambe le metodologie, infatti nessun vincolo riesce a migliorare l'AIC e il BIC. In entrambi i modelli i parametri sono tutti significativi.

MODELLO DI PARTENZA	
LISREL: AIC=40.92	LCA:AIC=2548 (BIC=2618)

Il modello migliore è quello in cui X influisce su V e K e quest'ultima su V, come rappresentato dal grafico Lisrel:

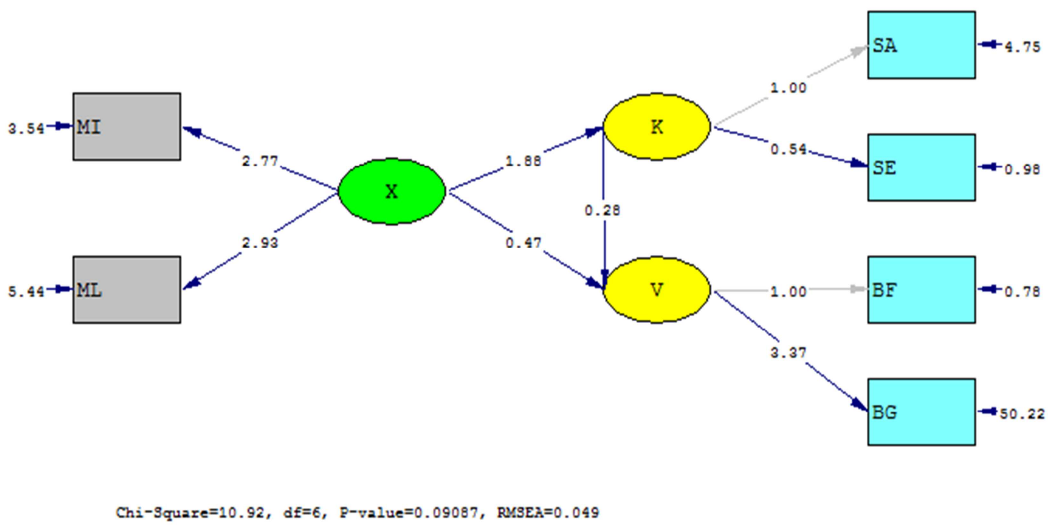


Figura 3: path diagram modello interesse-soddisfazione-comportamento futuro su prosciutto

Osservato che i modelli sono interpretabili in maniera uguale, si procede con l'analisi dei singoli parametri. Si è scelto di utilizzare la suddivisione dell'output di Lisrel (per tipo di regressore), cioè di analizzare i parametri λ, γ e β. Nell'output di Lem i parametri non sono divisi in questo modo bensì suddivisi per variabile risposta, quindi in un gruppo ci sono tutti i parametri che influiscono su K e in un altro gruppo tutti quelli che influiscono su V a prescindere che i regressori siano ξ o η.

- MATRICE LAMBDA

La matrice lambda rimane la stessa della matrice lambda della CFA poiché si sta stimando un modello completo e quindi X, K e Z hanno la stessa interpretazione:
LCA:
X: classe 1 molta int; classe 2 poca int
K: classe 1 molta sodd; classe 2 poca sodd
V: classe 1 comp molto pos; classe 2 comp poco pos

SEM:
X↑ =>int↑
K↑=> sodd↑
V↑=> comp pos↑

Ricordando quanto esposto nel capitolo 3 sul confronto dei parametri nelle 2 metodologie (SEM e LCA) in presenza di variabili latenti, risulta dalla matrice dei λ che le variabili latenti X, K e V sono variabili decrescenti in LCA e crescenti in SEM. Si ottiene lo stesso risultato con le 2 metodologie se i parametri strutturali hanno tutti lo stesso segno in quanto tutte le 3 variabili latenti in LCA sono di "segno" opposto a quelle in SEM.

- MATRICE GAMMA

Come indica la tabella seguente, in entrambi i modelli un aumento dell'interesse per il prosciutto di Sauris (quindi un aumento di X per SEM, un passaggio dalla classe 1 alla classe 2 per LCA), porta ad un aumento della soddisfazione (un aumento di K per SEM, un passaggio dalla classe 2 alla classe 1 per LCA) e ad un aumento del comportamento positivo (un aumento di V per SEM, un passaggio dalla classe 2 alla classe 1 per LCA).

L C A	X↑=>K↑	X.K effect lambda std err z-value exp(lambda) Wald df prob 1 1 0.6906 0.1270 5.440 1.9950 29.59 1 0.000 Ad un X=1 (molto interesse) è associata una più alta probabilità di essere in K=1 (molta sodd) e così ad X=2 (poco interesse) è associata una più alta probabilità di avere un K=2 (poca sodd)
	X↑=>V↑	X.V effect lambda std err z-value exp(lambda) Wald df prob 1 1 0.5105 0.1858 2.748 1.6662 7.55 1 0.006 Ad un X=1 (molto interesse) è associata una più alta probabilità di essere in V=1 (comp molto positivo) e così ad X=2 (poco interesse) è associata una più alta probabilità di avere un V=2 (comp poco pos)

L I S R E L	X↑=>K↑	GAMMA X -------- K 1.88 (0.21) 8.84
	X↑=>V↑	V 0.47 (0.13) 3.51

- MATRICE BETA

Come indica la tabella seguente, un aumento della soddisfazione (un aumento di K per SEM, un passaggio dalla classe 1 a quella 2 per LCA) porta ad un aumento del comportamento positivo futuro (un aumento di V per SEM, un passaggio dalla classe 1 a quella 2 per LCA).

<table>
<tr><td>L
C
A</td><td>K↑=>V↑</td><td>K.V
effect lambda std err z-value exp(lambda) Wald df prob
1 1 0.4999 0.1408 3.551 1.6485 12.61 1 0.000

Ad un K=1 (molta sodd) è associata una più alta probabilità di essere in V=1 (comp molto pos) e così ad K=2 (poca sodd) è associata una più alta probabilità di avere un V=2 (comp poco pos)</td></tr>
<tr><td>L
I
S
R
E
L</td><td>K↑=>V↑</td><td>BETA

K V
-------- --------
K - - - -
V 0.28 - -
(0.04)
6.81</td></tr>
</table>

- EFFETTI INDIRETTI E TOTALI

L’effetto totale dell’interesse per il prosciutto è positivo in entrambi i modelli, cioè un aumento della soddisfazione aumenta il comportamento positivo futuro, sia grazie all’effetto diretto che a quello indiretto.

L C A	X↑=>V↑	p(v|x) 1|1 0.7200 2|1 0.2800 1|2 0.1267 2|2 0.8733 Da questa tabella si possono calcolare α e β (valori che possono essere calcolati anche attraverso i parametri moltiplicativi). Già questa prima analisi indica che ad un X=1 (molto interesse) è associata una più alta probabilità di essere in V=1 (comp molto pos) e ad un X=2 (poco interesse) è associato un V=2 (comp poco pos), quindi risulta esserci un effetto totale positivo. In questo caso si è sicuri che l'effetto totale sarà maggiore dell'effetto diretto in quanto quest'ultimo e l'indiretto sono dello stesso segno, questo si può controllare calcolandolo direttamente: α=0.1450 β=0.3206 $\mu^{X=1,V=1} = \sqrt[4]{\frac{\alpha+1-\beta}{\alpha\beta}}$=2.0519 =>ln($\mu^{X=1,V=1}$) $= 0.7187$ L'effetto totale è di 0.718 che è maggiore di 0.5105 Per calcolare l'effetto indiretto totale si utilizzano anche gli effetti one-variabile : $\tau^{X=1} = 1.2048$ $\tau^{K=1} = 0.8927$ $\tau^{V=1} = 0.7890$ Effetto_ind_tot=1.2315 =>ln(Effetto_ind_tot)=0.2082 Effetto_cella=0.9074=> ln(Effetto_cella)=-0.097
L I S R E L	X↑=>V↑	Total and Indirect Effects Total Effects of KSI on ETA X -------- V 1.01 (0.11) 8.76 Indirect Effects of KSI on ETA X -------- V 0.53 (0.09) 5.78

4.4 Caso 3:modello solo effetto indiretto

In questo modello si cerca di capire come l'interesse verso il prosciutto possa influire sui comportamenti futuri riguardanti la festa di Sauris. Risulta dall'analisi del modello che all'aumentare dell'interesse, aumenta il desiderio di ritornare alla festa e di consigliarla ad altri. L'interesse influisce però sul comportamento futuro solo indirettamente, cioè attraverso la soddisfazione. I partecipanti più interessati al prosciutto risultano essere anche i più soddisfatti dalla festa e di conseguenza sono anche più propensi a ritornarci e a raccomandarla. L'interesse sul prosciutto fa aumentare la possibilità di un ritorno e di una pubblicità (intesa come passaparola) positiva della festa. Il modello con solo effetto indiretto non ha un buon adattamento ai dati, ma è stato comunque qui esposto perché mostra un esempio di come sia il modello SEM che quello causale categoriale a classi latenti portino alla scelta dello stesso modello. Le variabili osservate di questo sottomodello sono MI e ML, indicatori dell'interesse verso il prosciutto, SA e SE, indicatori della soddisfazione e BA e BE, indicatori del comportamento futuro verso la festa.

4.4.1 CFA

Per poter confrontare la metodologia LCA e quella SEM bisogna scrivere lo stesso modello con le 2 metodologie:

SEM

$$\begin{bmatrix} MI \\ ML \\ SA \\ SE \\ BA \\ BE \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ \lambda_2 & 0 & 0 \\ 0 & \lambda_3 & 0 \\ 0 & \lambda_4 & 0 \\ 0 & 0 & \lambda_5 \\ 0 & 0 & \lambda_6 \end{bmatrix} \begin{bmatrix} X \\ K \\ H \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix}$$

$$\text{TD}=\begin{bmatrix} \theta_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & \theta_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & \theta_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \theta_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & \theta_{66} \end{bmatrix} \qquad \text{PH}=\begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{12} & \phi_{22} & \phi_{23} \\ \phi_{13} & \phi_{23} & \phi_{33} \end{bmatrix}$$

LCA

$$\pi^{X=x\ K=k\ H=h\ MI=mi\ ML=ml\ SA=sa\ SE=se\ BA=ba\ BE=be}$$
$$= \pi^{X=x\ K=k\ H=h} \pi^{MI=mi|X=x} \pi^{ML=ml|X=x} \pi^{SA=sa|K=k} \pi^{SE=se|K=k} \pi^{BA=ba|H=h} \pi^{BE=be|H=h}$$

- MATRICE LX

<table>
<tr><td rowspan="3">L
C
A</td><td>X</td><td>1=
POCO
INT
2=
MOLTO
INT</td><td>* TABLE X.MI [or P(MI|X)] *
effect lambda std err z-value exp(lambda) Wald df prob
MI
1 -0.3604 0.1387 -2.599 0.6974 6.76 1 0.009
X.MI
1 1 0.7316 0.1254 5.835 2.0783 34.04 1 0.000
* TABLE X.ML [or P(ML|X)] *
effect lambda std err z-value exp(lambda) Wald df prob
ML
1 -0.2547 0.1764 -1.444 0.7752 2.08 1 0.149
X.ML
1 1 0.8493 0.1589 5.345 2.3379 28.57 1 0.000</td></tr>
<tr><td>K</td><td>1=
MOLTA
SOD
2=
POCA
SOD</td><td>* TABLE K.SA [or P(SA|K)] *
effect lambda std err z-value exp(lambda) Wald df prob
SA
1 -0.1378 0.1856 -0.742 0.8713 0.55 1 0.458
K.SA
1 1 -1.0720 0.1598 -6.708 0.3423 45.00 1 0.000
* TABLE K.SE [or P(SE|K)] *
effect lambda std err z-value exp(lambda) Wald df prob
SE
1 -0.0237 0.1262 -0.187 0.9766 0.04 1 0.851
K.SE
1 1 -0.8813 0.1092 -8.068 0.4142 65.09 1 0.000</td></tr>
<tr><td>H</td><td>1=
COMP
POCO
POS
2=
COMP
MOLTO
POS</td><td>* TABLE H.BA [or P(BA|H)] *
effect lambda std err z-value exp(lambda) Wald df prob
BA
1 -0.0608 0.0987 -0.616 0.9410 0.38 1 0.538
H.BA
1 1 0.5725 0.0914 6.265 1.7728 39.25 1 0.000
* TABLE H.BE [or P(BE|H)] *
effect lambda std err z-value exp(lambda) Wald df prob
BE
1 -1.0583 0.6930 -1.527 0.3470 2.33 1 0.127
H.BE
1 1 1.1162 0.6818 1.637 3.0531 2.68 1 0.102</td></tr>
<tr><td rowspan="3">L
I
S
R
E
L</td><td>X</td><td>↑ INT</td><td>LAMBDA-X
K H X
-------- -------- --------
MI -- -- 2.73
(0.19)
14.18
ML -- -- 2.97
(0.21)
13.86</td></tr>
<tr><td>K</td><td>↑
SODD</td><td>LAMBDA-X
K H X
-------- -------- --------
SA 1.00 -- --
SE 0.52 -- --
(0.03)
15.24</td></tr>
<tr><td>H</td><td>↑
COMP
FESTA</td><td>LAMBDA-X
K H X
-------- -------- --------
BA -- 1.00 --
BE -- 0.90 --
(0.09)
9.81</td></tr>
</table>

Nel modello SEM X, K e V sono variabili crescenti, in LCA X e H sono crescenti e K è decrescente

- MATRICE PHI

Come già spiegato, mentre in SEM i parametri sono effettivamente le varianze – covarianze delle variabili, in LCA ciò non avviene.

LCA	X	K	X.K effect lambda std err z-value exp(lambda) Wald df prob 1 1 -0.3716 0.2256 -1.648 0.6896 2.71 1 0.099
		H	X.H effect lambda std err z-value exp(lambda) Wald df prob 1 1 0.4241 0.2545 1.666 1.5283 2.78 1 0.096
	K	H	K.H effect lambda std err z-value exp(lambda) Wald df prob 1 1 -1.0149 0.3669 -2.766 0.3624 7.65 1 0.006
LISREL	X	X	1
		K	1.92 (0.21) 9.09
		H	0.89 (0.14) 6.48
	K	K	10.80 (1.23) 8.76
		H	4.69 (0.59) 7.98
	H	H	3.27 (0.53) 6.19

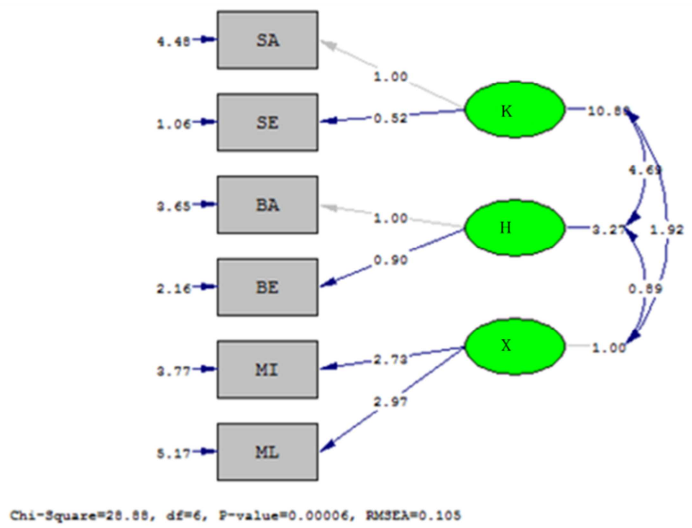


Figura 3: path diagram CFA modello interesse-soddisfazione-comportamento futuro sulla festa

4.4.2 Analisi strutturale

Si parte da un modello completo, in cui ci siano tutti i β e tutti i γ: questo non è un buon modello e come si vedrà è migliorabile. Il modello completo è nuovamente rappresentato nella metodologia LCA e in quella SEM come segue:

SEM

$$\begin{bmatrix} K \\ H \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} \mathrm{X} + \begin{bmatrix} 0 & 0 \\ \beta_{2\,1} & 0 \end{bmatrix} \begin{bmatrix} K \\ H \end{bmatrix} + \begin{bmatrix} \varsigma_1 \\ \varsigma_2 \end{bmatrix} \quad \mathrm{PSI}=\begin{bmatrix} \psi_{11} & 0 \\ 0 & \psi_{22} \end{bmatrix} \quad \mathrm{PH}=[\phi_{11}]$$

LCA

$$\pi^{X=x\ K=k\ H=h} = \pi^{X=x} \pi^{K=k|X=x} \pi^{H=h|X=x\ K=k}$$

Il modello SEM, come il modello LCA, ha il parametro di XH non significativo e per questo si prova a porlo uguale a 0. Il modello ristretto è rappresentato nella metodologia LCA e in quella SEM come segue:

SEM

$$\begin{bmatrix} K \\ H \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ 0 \end{bmatrix} \mathrm{X} + \begin{bmatrix} 0 & 0 \\ \beta_{2\,1} & 0 \end{bmatrix} \begin{bmatrix} K \\ H \end{bmatrix} + \begin{bmatrix} \varsigma_1 \\ \varsigma_2 \end{bmatrix} \quad \mathrm{PSI}=\begin{bmatrix} \psi_{11} & 0 \\ 0 & \psi_{22} \end{bmatrix} \quad \mathrm{PH}=[\phi_{11}]$$

LCA

$$\pi^{X=x\ K=k\ H=h} = \pi^{X=x} \pi^{K=k|X=x} \pi^{H=h|K=k}$$

Il modello di partenza, cioè quello completo, è rappresentato in forma Lisrel dal grafico sottostante, in cui si osserva che le frecce (relazioni) sono tutte presenti:

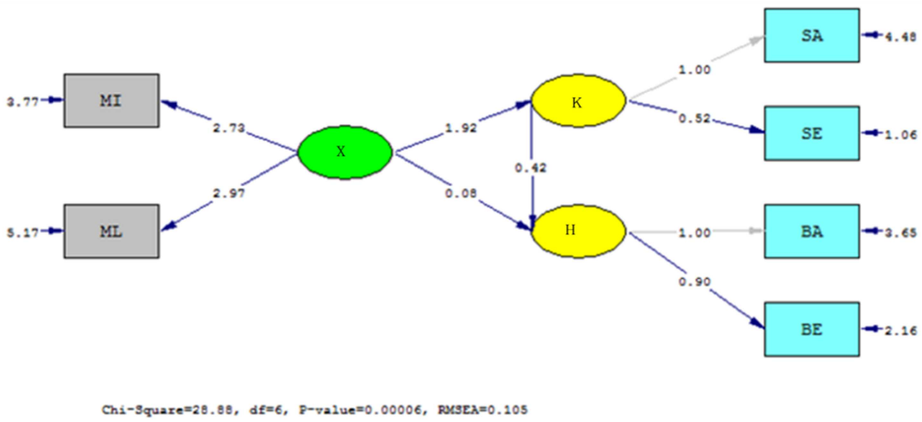


Figura 4:path diagram modello completo interesse-soddisfazione-comportamento futuro sulla festa

I valori esposti nella tabella 16 indicano che il parametro γ^{XV}è non significativo, infatti ipotizzando un test a livello 0.05:

SEM $H_0: \hat{\gamma}^{XV} = 0 \; vs \; H_1: \hat{\gamma}^{XV} \neq 0$

LCA $H_0: \underline{\hat{\gamma}}^{X=x\,V=v} = 0 \; vs \; H_1: \underline{\hat{\gamma}}^{X=x\,V=V} \neq 0$ con *x*, *v*=1,2

La regione di accettazione del test diviene

$$-1.96 \leq \frac{\hat{\gamma}^{XV}}{se(\hat{\gamma}^{XV})} \leq 1.96$$

Per il modello SEM il valore della statistica t è 0.57, per il modello LCA è 1.66, cioè entrambe le statistiche sono interne all'intervallo e quindi si accetta l'ipotesi che questo parametro sia uguale a 0

Tabella 16: parametri modello completo nelle 2 metodologie

SEM

```
BETA                                          GAMMA
                                                     X
        K     H                                   --------
     --------  --------                      K     1.92
  K     - -      - -                               (0.21)
  H    0.42       - -                               9.09
       (0.05)                                H      0.08
        8.04                                       (0.15)
                                                    0.57
```

LCA

```
* TABLE X [or P(X)] *
  effect          lambda  std err  z-value   exp(lambda)    Wald  df  prob
  X
   1              0.0550   0.1183    0.465      1.0565      0.22   1 0.642
* TABLE X.K [or P(K|X)] *
  effect          lambda  std err  z-value   exp(lambda)    Wald  df  prob
  K
   1             -0.0632   0.1524   -0.415      0.9388      0.17   1 0.678
  X.K
   1 1           -0.6660   0.1141   -5.837      0.5138     34.07   1 0.000
* TABLE X.K.H [or P(H|X.K)] *
  effect          lambda  std err  z-value   exp(lambda)    Wald  df  prob
  H
   1              0.4299   0.4411    0.975      1.5371      0.95   1 0.330
  X.H
   1 1            0.4241   0.2545    1.666      1.5282      2.78   1 0.096
  K.H
   1 1           -1.0149   0.3671   -2.765      0.3624      7.64   1 0.006
```

Il modello migliorato (per esempio nel modello log-lineare il BIC si riduce da 2592 a 2589) è quello rappresentato in forma Lisrel nel diagramma illustrato qui sotto. X influisce direttamente solo su K e solo indirettamente su H:

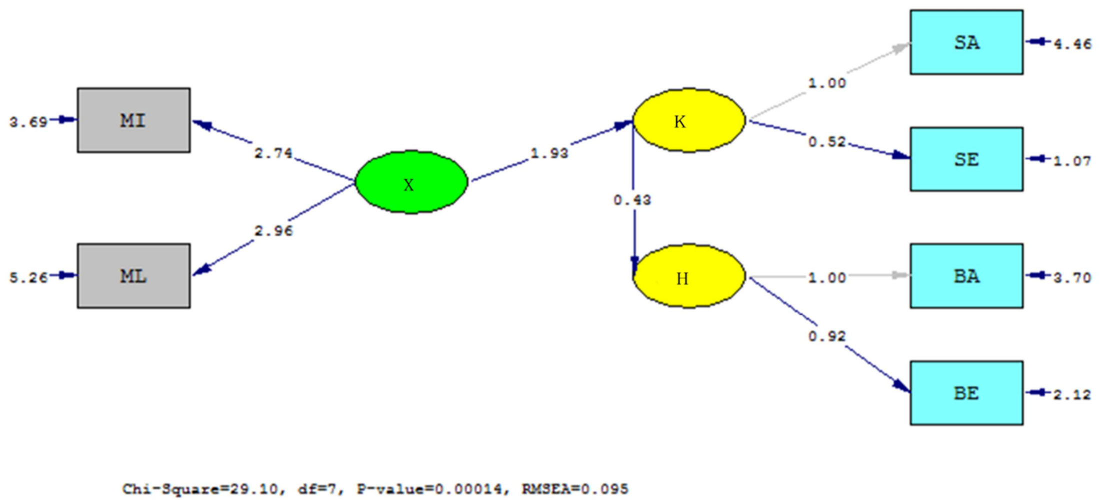


Figura 5:path diagram modello ristretto interesse-soddisfazione-comportamento futuro sulla festa

Come nel caso del modello completo (cioè il modello in cui l'interesse sul prosciutto influisce direttamente e indirettamente sul comportamento futuro riguardante il medesimo prodotto) si analizzano i parametri secondo la distinzione Lisrel.

- LAMBDA

La matrice Lambda varia rispetto alla medesima matrice del CFA in quanto questo è non è un modello completo; infatti le 2 matrici lambda sono uguali solo nel caso di un modello completo.

<table>
<tr>
<td rowspan="3">L
C
A</td>
<td>X</td>
<td>1=
POCO
INT
2=
MOLTO
INT</td>
<td>* TABLE X.MI [or P(MI|X)] *
effect lambda std err z-value exp(lambda) Wald df prob
MI
1 -0.3982 0.1661 -2.398 0.6715 5.75 1 0.016
X.MI
1 1 0.7883 0.1475 5.346 2.1997 28.58 1 0.000
* TABLE X.ML [or P(ML|X)] *
effect lambda std err z-value exp(lambda) Wald df prob
ML
1 -0.2487 0.1645 -1.512 0.7798 2.29 1 0.130
X.ML
1 1 0.8008 0.1445 5.540 2.2273 30.69 1 0.000

MI e ML sono variabili dicotomiche che rappresentano rispettivamente “l’interesse nell’approfondire la conoscenza del prosciutto di Sauris” e “l’interesse alla possibilità di assaggiare il prodotto”. La classe 1 significa poco interesse, la classe 2 molto interesse; poiché alla classe X=1 è associata in entrambi i casi la classe 1 e viceversa, ciò significa che la variabile latente X ha come valore 2 il molto interesse verso il prodotto e come valore 1 il poco interesse.</td>
</tr>
<tr>
<td>K</td>
<td>1=
MOLTA
SODD
2=
POCA
SODD</td>
<td>* TABLE K.SA [or P(SA|K)] *
effect lambda std err z-value exp(lambda) Wald df prob
SA
1 -0.1402 0.1823 -0.769 0.8691 0.59 1 0.442
K.SA
1 1 -1.0393 0.1468 -7.079 0.3537 50.12 1 0.000
* TABLE K.SE [or P(SE|K)] *
effect lambda std err z-value exp(lambda) Wald df prob
SE
1 -0.0312 0.1297 -0.241 0.9693 0.06 1 0.810
K.SE
1 1 -0.8736 0.1059 -8.249 0.4175 68.04 1 0.000

SA e SE sono variabili dicotomiche che rappresentano rispettivamente “pensare alla festa del prosciutto di Sauris mi rende felice” e “la festa del prosciutto di Sauris mi dà un senso di gioia”. La classe 1 significa poca felicità/gioia, la classe 2 molta; poiché alla classe K=1 è associata in entrambi i casi la classe 2 e viceversa, ciò significa che la variabile latente K ha come valore 1 la molta soddisfazione sulla festa, come valore 2 la poca.</td>
</tr>
<tr>
<td>H</td>
<td>1=
COMP
POCO
POS
2=
COMP
MOLTO
POS</td>
<td>* TABLE H.BA [or P(BA|H)] *
effect lambda std err z-value exp(lambda) Wald df prob
BA
1 -0.0544 0.1046 -0.520 0.9471 0.27 1 0.603
H.BA
1 1 0.5623 0.0904 6.217 1.7547 38.65 1 0.000
* TABLE H.BE [or P(BE|H)] *
effect lambda std err z-value exp(lambda) Wald df prob
BE
1 -1.4030 2.9823 -0.470 0.2459 0.22 1 0.638
H.BE
1 1 1.4788 2.9596 0.500 4.3877 0.25 1 0.617

BA e BE sono variabili dicotomiche che rappresentano rispettivamente “parlerò in modo positivo della festa del prosciutto di Sauris” e “verrò alla festa del prosciutto di Sauris o l’anno prossimo o l’anno successivo”. La classe 1 significa che non parlerà/non verrà alla festa, la classe 2 parlerà/verrà; poiché alla classe H=1 è associata in entrambi i casi la classe 1 e viceversa, ciò significa che la variabile latente H ha come valore 1 un comportamento negativo/poco positivo futuro verso la festa del prosciutto di Sauris, come valore 2 un comportamento positivo</td>
</tr>
</table>

L I S R E L	X	$\uparrow$ INT	LAMBDA-X X -------- MI 2.74 (0.19) 14.22 ML 2.96 (0.21) 13.78
	K	$\uparrow$ SODD	LAMBDA-Y K H -------- -------- SA 1.00 - - SE 0.52 - - (0.03) 15.38
	H	$\uparrow$ COMP	LAMBDA-Y K H -------- -------- BA - - 1.00 BE - - 0.92 (0.09) 9.77

Ricordando quanto esposto nel capitolo 3 sul confronto dei parametri nelle 2 metodologie (SEM e LCA) in presenza di variabili latenti, risulta dalla matrice dei λ che le variabili latenti X e H sono variabili crescenti sia in LCA che in SEM e la variabile latente K è decrescente in LCA e crescente in SEM. Si ottiene lo stesso risultato con le 2 metodologie se i parametri strutturali che spiegano il legame X$\rightarrow$H hanno tutti lo stesso segno (in quanto le 2 variabili latenti in LCA sono di "segno" uguale a quelle in SEM) e se i parametri strutturali che spiegano il legame X$\rightarrow$K e K$\rightarrow$H sono di segno opposto (in quanto le variabili latenti X e K in LCA sono di "segno" uguale a quelle in SEM e la variabile latente K è di segno opposto).

- MATRICE GAMMA

In entrambi i modelli un aumento dell'interesse per il prosciutto di Sauris (quindi un aumento di X per SEM, un passaggio dalla classe 1 alla classe 2 per LCA), porta ad un aumento della soddisfazione (un aumento di K per SEM, un passaggio dalla classe 2 alla classe 1 per LCA) .

LCA	X↑=>K↑	X.K effect lambda std err z-value exp(lambda) Wald df prob 1 1 -0.7037 0.1171 -6.012 0.4948 36.14 1 0.000 Ad un X=1 (poco interesse) è associata una più alta probabilità di essere in K=2 (poca sodd) e così ad X=2 (molto interesse) è associata una più alta probabilità di avere un K=1 (molta sodd)
LISREL	X↑=>K↑	GAMMA X -------- K 1.93 (0.21) 9.26 H - -

- MATRICE BETA

Un aumento della soddisfazione (un aumento di K per SEM, un passaggio dalla classe 2 a quella 1 per LCA) porta ad un aumento del comportamento positivo futuro (un aumento di H per SEM, un passaggio dalla classe 1 a quella 2 per LCA)

LCA	K↑=>H↑	K.H effect lambda std err z-value exp(lambda) Wald df prob 1 1 -1.1509 0.3770 -3.053 0.3163 9.32 1 0.002 Ad un K=1 (molta sodd) è associata una più alta probabilità di essere in H=2 (comp molta pos) e così ad K=2 (poca sodd) è associata una più alta probabilità di avere un H=1 (comp poco pos)
LISREL	K↑=>H↑	BETA K H -------- -------- K - - - - H 0.43 - - (0.04) 9.87

- EFFETTI INDIRETTI E TOTALI

L’effetto totale dell’interesse per il prosciutto è positivo in entrambi i modelli, cioè un aumento dell’interesse aumenta il comportamento positivo futuro con l’effetto indiretto, poiché un aumento dell’interesse fa aumentare la soddisfazione che a sua volta aumenta il comportamento positivo futuro.

L C A	X↑=>H↑	p(h|x) 1|1 0.8166 2|1 0.1834 1|2 0.3446 2|2 0.6554 Da questa tabella si possono calcolare α e β (valori che possono essere calcolati anche attraverso i parametri moltiplicativi). Già questa prima analisi indica che ad un X=1 (poco interesse) è associata una più alta probabilità di essere in H=1 (comp poco pos) e ad un X=2 (molto interesse) è associato un H=2 (comp molto pos), quindi c'è un effetto totale positivo. In questo caso l'effetto totale sarà uguale all'effetto indiretto α=0.5257 β=0.2799 $\mu^{X=1,V=1} = \sqrt[4]{\frac{\alpha+1-\beta}{\alpha\beta}}$=1.7057 =>ln($\mu^{X=1,V=1}$) = 0.5340 L'effetto totale è di 0.5340 $\tau^{X=1}$ =1.0678 $\tau^{K=1}$ = 0.9386 $\tau^{H=1}$ =1.4561 Effetto_ind_tot=Effetto_ind=1.7057=> ln(Effetto_ind)=0.5340 Effetto_cella=1 => ln(Effetto_cella)=0 Effetti totali = effetti indiretti In un modello senza l'effetto diretto di A su C, l'effetto totale è uguale all'effetto indiretto in quanto l'effetto cella è sempre uguale a 1 in termini moltiplicativi e uguale a 0 in termini log-lineari
L I S R E L	X↑=>H↑	Total and Indirect Effects Total Effects of KSI on ETA X -------- K 1.93 (0.21) 9.26 H 0.84 (0.11) 7.40 Indirect Effects of KSI on ETA X -------- K - - H 0.84 (0.11) 7.40

4.5 Caso 4: modello effetto cella

Concludo il capitolo con un modello completo che è particolarmente interessante per vedere il ruolo dell'effetto cella, "interazione" insita solo nei modelli log-lineari.
In questo modello si cerca di capire come la "qualità" della festa di Sauris possa influire sui comportamenti futuri riguardanti il prodotto sponsorizzato durante la manifestazione. Si vede dall'analisi del modello che all'aumentare della qualità, aumenta il desiderio di riacquistare il prosciutto e di consigliarlo ad altri. La qualità influisce in 2 modi sul comportamento futuro: sia direttamente che indirettamente, cioè attraverso la soddisfazione. I partecipanti che hanno apprezzato maggiormente la festa risultano essere anche i più soddisfatti dalla festa e i più propensi a riacquistare il prodotto ed a raccomandarne l'acquisto ad altri. La festa, quindi, fa aumentare la propensione all'acquisto del prodotto e la sua pubblicizzazione (intesa come passaparola). Le variabili osservate di questo sottomodello sono QG e QV, indicatori della qualità della festa, SA e SE, indicatori della soddisfazione e BF e BG, indicatori del comportamento futuro verso il prosciutto.

4.5.1 CFA

Per poter confrontare la metodologia LCA con quella SEM si scrive lo stesso modello con le 2 metodiche:

SEM

$$\begin{bmatrix} QG \\ QV \\ SA \\ SE \\ BF \\ BG \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ \lambda_2 & 0 & 0 \\ 0 & \lambda_3 & 0 \\ 0 & \lambda_4 & 0 \\ 0 & 0 & \lambda_5 \\ 0 & 0 & \lambda_6 \end{bmatrix} \begin{bmatrix} W \\ K \\ V \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix}$$

$$\text{TD}=\begin{bmatrix} \theta_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & \theta_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & \theta_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \theta_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & \theta_{66} \end{bmatrix} \quad \text{PH}=\begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{12} & \phi_{22} & \phi_{23} \\ \phi_{13} & \phi_{23} & \phi_{33} \end{bmatrix}$$

LCA

$$\pi^{W=w\ K=k\ V=v\ QA=qa\ QV=qv\ SA=sa\ SE=se\ BF=bf\ BG=bg} = \pi^{W=w\ K=k\ V=v} \pi^{QA=qa|W=w} \pi^{QV=qv|W=w} \pi^{SA=sa|K=k} \pi^{SE=se|K=k} \pi^{BF=bf|V=v} \pi^{BG=bg|V=v}$$

- MATRICE LX

<table>
<tr><td rowspan="3">L
C
A</td><td>W</td><td>1=
POCA
QUAL.
2=
MOLTA
QUAL.</td><td>* TABLE W.QG [or P(QG|W)] *
effect lambda std err z-value exp(lambda) Wald df prob
QG
1 0.4765 0.1692 2.816 1.6105 7.93 1 0.005
W.QG
1 1 0.8456 0.1529 5.531 2.3295 30.59 1 0.000
* TABLE W.QV [or P(QV|W)] *
effect lambda std err z-value exp(lambda) Wald df prob
QV
1 -0.5987 0.1585 -3.776 0.5495 14.26 1 0.000
W.QV
1 1 0.7606 0.1467 5.184 2.1395 26.87 1 0.000

QG e QV sono variabili dicotomiche che rappresentano rispettivamente “qualità festa del prosciutto di Sauris” e “complessivamente il livello dei servizi offerti al pubblico alla festa del prosciutto di Sauris”. La classe 1 significa poca qualità, la classe 2 molta qualità; poiché alla classe W=1 è associata in entrambi i casi la classe 1 e viceversa, ciò significa che la variabile latente W ha come valore 1 poca qualità della festa e come valore 2 la molta qualità.</td></tr>
<tr><td>K</td><td>1=
POCA
SOD
2=
MOLTA
SOD</td><td>* TABLE K.SA [or P(SA|K)] *
effect lambda std err z-value exp(lambda) Wald df prob
SA
1 -0.3380 0.1727 -1.958 0.7132 3.83 1 0.050
K.SA
1 1 1.0116 0.1610 6.285 2.7501 39.50 1 0.000
* TABLE K.SE [or P(SE|K)] *
effect lambda std err z-value exp(lambda) Wald df prob
SE
1 -0.3491 0.2401 -1.454 0.7053 2.11 1 0.146
K.SE
1 1 1.1534 0.2250 5.126 3.1690 26.28 1 0.000

SA e SE sono variabili dicotomiche che rappresentano rispettivamente “pensare alla festa del prosciutto di Sauris mi rende felice” e “la festa del prosciutto di Sauris mi dà un senso di gioia”. La classe 1 significa poca felicità/gioia, la classe 2 molta; poiché alla classe K=1 è associata in entrambi i casi la classe 1 e viceversa, ciò significa che la variabile latente K ha come valore 1 la poca soddisfazione sulla festa, come valore 2 la molta.</td></tr>
<tr><td>V</td><td>1=
MOLTA
2=
POCA</td><td>* TABLE V.BF [or P(BF|V)] *
effect lambda std err z-value exp(lambda) Wald df prob
BF
1 -0.1323 0.0926 -1.429 0.8761 2.04 1 0.153
V.BF
1 1 -0.5881 0.0938 -6.271 0.5554 39.33 1 0.000
* TABLE V.BG [or P(BG|V)] *
effect lambda std err z-value exp(lambda) Wald df prob
BG
1 -1.1254 0.6218 -1.810 0.3245 3.28 1 0.070
V.BG
1 1 -1.1189 0.6228 -1.796 0.3266 3.23 1 0.072

BF e BG sono variabili dicotomiche che rappresentano rispettivamente “parlerò in modo positivo del prosciutto di Sauris” e “acquisterò il prosciutto di Sauris più spesso”. La classe 1 significa che non si parlerà/non si comprerà , la classe 2 che si comprerà/si parlerà; poiché alla classe V=1 è associata in entrambi i casi la classe 2 e viceversa, ciò significa che la variabile latente V ha come valore 1 un comportamento futuro positivo nei confronti del prosciutto di Sauris, come valore 2 un comportamento negativo o meno positivo.</td></tr>
</table>

LISREL	X	↑ QUAL	LAMBDA-X W K V -------- -------- -------- QG 0.81 - - - - (0.06) 13.49 QV 0.74 - - - - (0.06) 12.57
	K	↑ SODD	LAMBDA-X W K V -------- -------- -------- SA - - 1.00 - - SE - - 1.09 - - (0.07) 14.72
	V	↑ COMP	LAMBDA-X W K V -------- -------- -------- BA - - - - 1.00 BE - - - - 1.05 (0.10) 10.28

- MATRICE PHI

			effect	lambda	std err	z-value	exp(lambda)	Wald	df	prob
LCA	W	W	W 1	-0.2922	0.3043	-0.960	0.7466	0.92	1	0.337
		K	W.K 1 1	0.0110	0.4938	0.022	1.0110	0.00	1	0.982
		V	W.V 1 1	-0.8871	0.5834	-1.521	0.4118	2.31	1	0.128
	K	K	K 1	-0.1184	0.3073	-0.385	0.8883	0.15	1	0.700
		V	K.V 1 1	-1.1933	0.5315	-2.245	0.3032	5.04	1	0.025
	V	V	V 1	-0.5876	0.4432	-1.326	0.5557	1.76	1	0.185

L I S R E L	W	W	1
		K	0.45 (0.05) 8.30
		V	0.40 (0.05) 7.48
	K	K	0.66 (0.08) 8.35
		V	0.45 (0.06) 7.99
	V	V	0.49 (0.08) 6.42

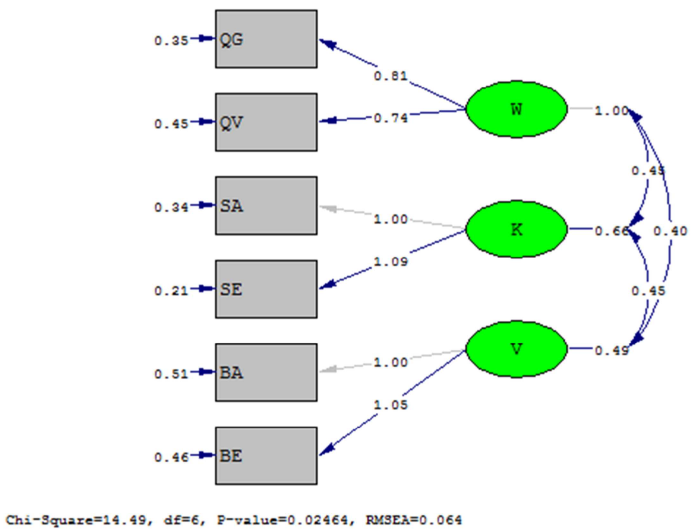


Figura 6: path diagram CFA modello qualità-soddisfazione-comportamento futuro su prosciutto

4.5.2 Analisi strutturale

Si parte da un modello completo, nel quale tutti i γ e tutti i β, in termini di Lisrel, sono compresi. Come per il modello SEM, anche per il modello loglineare causale a classi latenti gli indici di bontà del modello CFA e di quello completo rimangono gli stessi, questo perché porre una correlazione o un effetto diretto non cambia la bontà del modello. Per poter confrontare la metodologia LCA con quella SEM si scrive lo stesso modello con le 2 metodiche:

SEM

$$\begin{bmatrix} K \\ V \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} \mathrm{X} + \begin{bmatrix} 0 & 0 \\ \beta_{2\,1} & 0 \end{bmatrix} \begin{bmatrix} K \\ V \end{bmatrix} + \begin{bmatrix} \varsigma_1 \\ \varsigma_2 \end{bmatrix} \quad \mathrm{PSI}=\begin{bmatrix} \psi_{11} & 0 \\ 0 & \psi_{22} \end{bmatrix} \quad \mathrm{PH}=[\phi_{11}]$$

LCA

$$\pi^{W=w\ K=k\ V=v} = \pi^{W=w} \pi^{K=k|W=w} \pi^{V=v|W=w\ K=k}$$

In questo caso il modello completo è il miglior modello con entrambe le metodologie, infatti nessun vincolo riesce a migliorare l'AIC e il BIC. In entrambi i modelli i parametri sono tutti significativi.

MODELLO DI PARTENZA	
LISREL: AIC=44.49	LCA:AIC=2492 (BIC=2561)

Il modello migliore è quello in cui W influisce su V e K e quest'ultima su V, come rappresentato dal grafico Lisrel:

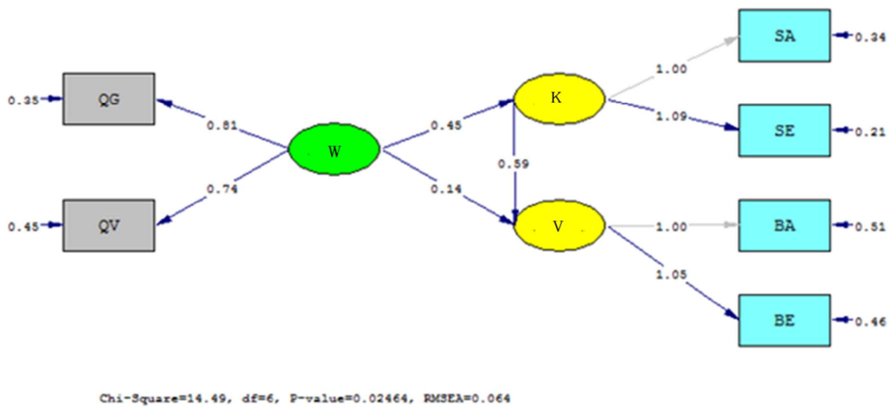


Figura 7:path diagram qualità-soddisfazione-comportamento futuro su prosciutto

Osservato che i modelli sono interpretabili in maniera uguale, si procede con l'analisi dei singoli parametri. Si è scelto di utilizzare la suddivisione dell'output di Lisrel (per tipo di regressore), cioè di analizzare i parametri λ, γ e β. Nell'output di Lem i parametri non sono divisi in questo modo ma per variabile risposta, quindi in un gruppo ci sono tutti i parametri che influiscono su K e in un altro gruppo tutti quelli che influiscono su V a prescindere che i regressori siano ξ o η.

- LAMBDA

La matrice lambda rimane la stessa della matrice lambda della CFA poiché si sta stimando un modello completo e quindi X, K e Z hanno la stessa interpretazione:
LCA:
W: classe1 poca qualità; classe 2 molta qualità
K: classe 1 poca sodd; classe 2 molta sodd
V: classe 1 comp molto pos; classe 2 comp poco pos
SEM:
W↑ =>qualità↑
K↑=> sodd↑
V↑=> comp pos↑

Ricordando quanto esposto nel capitolo 3 sul confronto dei parametri nelle 2 metodologie (SEM e LCA) in presenza di variabili latenti, risulta dalla matrice dei λ che le variabili latenti W e K sono variabili decrescenti, la variabile V è decrescente in LCA mentre in SEM sono tutte variabili crescenti. Si ottiene lo stesso risultato con le 2 metodologie se i parametri strutturali hanno tutti lo stesso segno in quanto tutte le 3 variabili latenti in LCA sono di "segno" opposto a quelle in SEM.
Si ottiene lo stesso risultato con le 2 metodologie se i parametri strutturali che spiegano il legame W→K hanno tutti lo stesso segno (in quanto le 2 variabili latenti in LCA sono di entrambe di "segno" opposto a quelle in SEM) e se i parametri strutturali che spiegano il legame W→V e K→V sono di segno opposto (in quanto le variabili latenti X e K in LCA sono di "segno" opposto a quelle in SEM e la variabile latente V è di segno uguale).

- MATRICE GAMMA

GAMMA rappresenta la matrice degli effetti diretti di W su K e su V nella metodologia SEM e in quella LCA.
Come indica la tabella seguente, in entrambi i modelli un aumento della qualità percepita della festa di Sauris (quindi un aumento di W per SEM, un passaggio dalla classe 1 alla classe 2 per LCA), porta ad un aumento della soddisfazione (un aumento di K per SEM, un passaggio dalla classe 1 alla classe 2 per LCA) e ad un aumento del comportamento positivo (un aumento di V per SEM, un passaggio dalla classe 2 alla classe 1 per LCA).

LCA	W↑=>K↑	W.K effect lambda std err z-value exp(lambda) Wald df prob 1 1 0.6193 0.1192 5.195 1.8577 26.99 1 0.000 Ad un W=1 (poca qualità) è associata una più alta probabilità di essere in K=1 (poca sodd) e così ad W=2 (molta qualità) è associata una più alta probabilità di avere un K=2 (molta sodd)
	W↑=>V↑	W.V effect lambda std err z-value exp(lambda) Wald df prob 1 1 -0.8871 0.5824 -1.523 0.4119 2.32 1 0.128 Ad un W=1 (poca qualità) è associata una più alta probabilità di essere in V=2 (poca) e così ad W=2 (molta qualità) è associata una più alta probabilità di avere un V=1 (molta)
LISREL	W↑=>K↑ W↑=>V↑	GAMMA W -------- K 0.45 (0.05) 8.30 V 0.14 (0.05) 2.50

- MATRICE BETA

BETA rappresenta la matrice degli effetti diretti di K su V nella metodologia SEM e in quella LCA.
Come indica la tabella seguente, un aumento della soddisfazione (un aumento di K per SEM, un passaggio dalla classe 1 a quella 2 per LCA) porta ad un aumento del comportamento positivo futuro (un aumento di V per SEM, un passaggio dalla classe 2 a quella 1 per LCA)

<table>
<tr><td>L
C
A</td><td>K↑=>V↑</td><td>

```
K.V
effect     lambda   std err   z-value   exp(lambda) Wald   df   prob
1 1       -1.1934    0.5307    -2.249      0.3032    5.06   1 0.025
```

Ad un K=1 (poca sodd) è associata una più alta probabilità di essere in V=2 (poca) e così ad K=2 (molta sodd) è associata una più alta probabilità di avere un V=1 (molta)</td></tr>
<tr><td>L
I
S
R
E
L</td><td>K↑=>V↑</td><td>

```
BETA

          K         V
       --------  --------
K         - -       - -
V        0.59       - -
        (0.07)
         7.94
```

</td></tr>
</table>

- EFFETTI TOTALI E INDIRETTI TOTALI

EFFETTI TOTALI rappresentano l'effetto che ha W su V sia direttamente che indirettamente attraverso K

EFFETTI INDIRETTI TOTALI rappresentano l'effetto che ha W su V solo attraverso K. L'effetto indiretto totale nella metodologia LCA si suddivide in effetto cella e in effetto indiretto. L'effetto indiretto è l'effetto che ha W su V attraverso K, l'effetto cella è l'effetto interazione dovuto alla presenza degli effetti diretti di W su V e di K su V (cioè l'effetto cella è l'influenza dell'effetto diretto W→V su l'effetto diretto K→V).
L'effetto totale dell'interesse per il prosciutto è positivo in entrambi i modelli, cioè un aumento della soddisfazione aumenta il comportamento positivo futuro, sia grazie all'effetto diretto che a quello indiretto.

L C A	W↑=>V↑	p(v|w) 1|1 0.0589 2|1 0.9411 1|2 0.6928 2|2 0.3072 Da questa tabella si possono calcolare α e β (valori che possono essere calcolati anche attraverso i parametri moltiplicativi). Questa prima analisi indica che ad un W=1 (poca qualità) è associata una più alta probabilità di essere in V=2 (poco) e a un W=2 (molta qualità) è associato un V=1 (molta), quindi questo mostra un effetto totale positivo. Si è sicuri che l'effetto totale è maggiore dell'effetto diretto in quanto quest'ultimo e l'indiretto sono dello stesso segno, questo si può controllare calcolandolo direttamente: α=2.2553 β=3.0634 $\mu^{X=1,V=1} = \sqrt[4]{\frac{\alpha+1-\beta}{\alpha\beta}}$=0.4082 =>ln$(\mu^{X=1,V=1}) = -0.8960$ L'effetto totale è di -0.8960 che è maggiore in valore assoluto di -0.8871 Per calcolare l'effetto indiretto totale si utilizzano anche gli effetti one-variabile : $\tau^{W=1}$=1.0447 $\tau^{K=1} = 1.2757$ $\tau^{V=1} = 0.5557$ Effetto_ind_tot=0.9911 => ln(Effetto_ind_tot)= -0.0089 Effetto_cella=1.6234 =>ln(Effetto_cella)=0.4845 Effetto_ind=0.6105=>ln(Effetto_ind)= -0.4934

L I S R E L	W↑=>V↑	Total and Indirect Effects Total Effects of KSI on ETA W -------- K 0.45 (0.05) 8.30 V 0.40 (0.05) 7.48 Indirect Effects of KSI on ETA W -------- K - - V 0.26 (0.04) 6.12

L'effetto totale di W su V produce una relazione interpretazionale positiva nelle 2 metodologie, ma l'effetto cella attenua l'effetto indiretto portando il valore numerico dell'effetto totale ad essere molto simile a quello dell'effetto diretto.

4.6 Conclusioni

Questo mi ha confermato quanto già intuito quando ho iniziato il confronto: i modelli SEM e quelli log-lineari causali mi forniscono un medesimo modello causale, ma l'intensità delle relazioni causali non sono confrontabili. Ritengo logica questa affermazione, in quanto non si possono confrontare numericamente "cose" diverse, infatti i parametri log-lineari misurano l'influenza di un particolare fattore sulla probabilità di essere nella classe i della variabile A, quelli SEM misurano l'influenza del fattore proprio sulla classe i della variabile A.

Appendice A

A1 Il significato di β e α

Come spiegato nel capitolo 1 (1.5) l'effetto totale μ^{CA} è funzione di β e α, ma cosa rappresentano β e α in concreto? In letteratura, come nel paper di Rao, Li e Roth (2008), si afferma che gli effetti diretti e totali in un modello con variabili dicotomiche si calcolano nel seguente modo:

effetto_direttoRLR=$log\left\{\frac{P(C=2|A=2,B)}{P(C=2|A=1,B)}\right\}$

effetto_totaleRLR=$log\left\{\frac{P(C=2|A=2)}{P(C=2|A=1)}\right\}$

Si sono chiamati effetti direttiRLR e totaliRLR per distinguerli da quelli effettivamente presenti nel capitolo. L'effetto β è la trasformazione esponenziale dell'inverso dell'effetto totaleRLR. Queste formule sono appropriate nei modelli che possono essere ricondotti ad un logit, ma perché nel modello causale non si possono utilizzare queste formule? L'esempio più semplice è quello dell'effetto diretto, che in un modello causale è rappresentato dalla variazione di τ^{CA}:

effetto_direttoRLR=$log\left\{\frac{\eta^{C|A=1,B}\tau^{C=2\ A=1}\tau^{C=2}\tau^{C=2\ B=2}}{\eta^{C|A=2,B}\tau^{C=2\ A=2}\tau^{C=2}\tau^{C=2\ B=2}}\right\} = log\left\{\frac{\eta^{C|A=1\ B}\tau^{C=2\ A=1}}{\eta^{C|A=2\ B}\tau^{C=2\ A=2}}\right\} \neq$
$log\left\{\frac{\tau^{C=2\ A=1}}{\tau^{C=2\ A=2}}\right\}$

Come si vede dalla formula questa differenza è dovuta alle costanti di normalizzazione η. Lo stesso problema si ha nell'effetto totale; ma mentre nell'effetto direttoRLR si riesce a suddividere l'effetto diretto da quello dovuto alle η, nell'effetto totale questo non si riesce a farlo. Qui entra in gioco α che serve per riuscire ad eliminare l'effetto delle η dall'effetto totale. L'effetto α si costruisce:

α=$\frac{P(C=1|A=2)}{P(C=2|A=2)}$= variazione di C

In un modello come quello log-lineare, la variazione di C fa variare sia τ^{C} che τ^{CA}, cioè:

$$\tau^{C=1} \neq \tau^{C=2}$$

$$\tau^{C=1\ A=2} \neq \tau^{C=2\ A=2}$$

Come fa l'effetto α ad eliminare dall'effetto totale il problema delle η? Si considera solo P(C|A) per calcolare l'effetto totale, come si vede nella formula dell'effetto totaleRLR:

$\pi^{C|A}$=$\eta^{C|A}\mu^{C\ A}\mu^{C}$ => dalla probabilità condizionata di C|A si ricava l'effetto totale di A su C

che non deve essere confusa con la condizionata utilizzata per stimare il modello che è data dall'identità:

$\pi^{C|AB}\equiv\ \eta^{C|AB}\tau^{C\ A}\tau^{C}\tau^{CB}$ => dalla probabilità condizionata C|AB si ricava l'effetto diretto di A su C

Dalla formula della condizionata C|A si trovano i valori di α e β

$\alpha=\frac{\eta^{C|A=1}\mu^{C=1\ A=1}\mu^{C=1}}{\eta^{C|A=1}\mu^{C=2\ A=2}\mu^{C=2}}=\frac{\mu^{C=1\ A=1}\mu^{C=1}}{\mu^{C=2\ A=2}\mu^{C=2}}$

$\beta=\frac{\eta^{C|A=1}\mu^{C=1\ A=1}\mu^{C=1}}{\eta^{C|A=2}\mu^{C=1\ A=2}\mu^{C=1}}=\frac{\eta^{C|A=1}\mu^{C=2\ A=1}}{\eta^{C|A=2}\mu^{C=2\ A=2}}$

Per verificarne la coerenza si effettua sia il calcolo dell'effetto totale in dummy code che in effect code.

- Utilizzando il dummy code i parametri diversi da 1 sono $\mu^{C=2\ A=2}$ e $\mu^{C=2}$, perciò il passaggio da A=1 a A=2 influenza C attraverso il rapporto $\mu^{C=2\ A=2}/1$,

quindi l'effetto totale è esattamente $\mu^{C=2\ A=2}$ o la sua trasformazione logaritmica, ma come si può scomporre $\mu^{C=2\ A=2}$?

Ricordando che

$\alpha=\frac{1}{\mu^{C=2\ A=2}\mu^{C=2}}$ (1.1)

$\beta=\frac{\eta^{C|A=1}}{\eta^{C|A=2}\mu^{C=2\ A=2}}$ (1.2)

dove

$$\eta^{C|A=1}=\frac{1}{1+\mu^{C=2}}$$

e

$$\eta^{C|A=2}=\frac{1}{1+\mu^{C=2}\mu^{C=2\ A=2}}$$

Dalla (1.1) si ricava $\mu^{C=2}$ in funzione di α e $\mu^{C=2\ A=2}$:

$$\mu^{C=2}=\frac{1}{\alpha\mu^{C=2\ A=2}}$$

che viene inserita nella (1.2)

$$\beta = \frac{1}{\mu^{C=2\ A=2}} \left[\frac{1}{1+\mu^{C=2}}(1+\mu^{C=2}\mu^{C=2\ A=2})\right]$$
$$=\frac{1}{\mu^{A=2\ C=2}} \left[\frac{1}{1+\frac{1}{\alpha\mu^{C=2\ A=2}}}\left(1+\frac{1}{\alpha\mu^{C=2\ A=2}}\mu^{C=2\ A=2}\right)\right]=$$
$$=\frac{1}{\mu^{C=2\ A=2}}\left[\frac{1}{1+\frac{1}{\alpha\mu^{C=2\ A=2}}}\left(1+\frac{1}{\alpha}\right)\right]=$$
$$=\frac{1}{\mu^{C=2\ A=2}}\left[\frac{\alpha\mu^{C=2\ A=2}}{\alpha\mu^{C=2\ A=2}+1}\left(1+\frac{1}{\alpha}\right)\right]=\frac{1}{\mu^{C=2\ A=2}}\left[\frac{\alpha\mu^{C=2\ A=2}}{\alpha\mu^{C=2\ A=2}+1}\left(\frac{\alpha+1}{\alpha}\right)\right]=\frac{\alpha+1}{\alpha\mu^{C=2\ A=2}+1}$$

Da questa relazione si ricava l'effetto totale $\mu^{C=2\ A=2}$:

$$\beta(\alpha\mu^{C=2\ A=2}+1) = \alpha+1 => \mu^{C=2\ A=2} = \frac{\alpha+1-\beta}{\alpha\beta}$$

- Utilizzando l'effect code i parametri da considerare sono $\mu^{C=2\ A=2}$ e $\mu^{C=2}$ perchè gli altri sono solo loro trasformazioni:

$$\mu^{C=2\ A=2} = \mu^{C=1\ A=1} = \frac{1}{\mu^{C=2\ A=1}} = \frac{1}{\mu^{C=1\ A=2}}$$
$$\mu^{\ C=2} = \frac{1}{\mu^{C=1}}$$

perciò il passaggio da A=1 a A=2 influenza C attraverso l'effetto totale $\mu^{C=2\ A=2}/\mu^{C=2\ A=1}$,che è $(\mu^{C=2\ A=2})^2$o $(\mu^{C=2\ A=1})^{-2}$ in quanto l'effetto totale con parametrizzazione effect code, come appena dimostrato, è il quadrato di $\mu^{C=2\ A=2}$ e l'effetto diretto è il quadrato di $\tau^{C=2A=2}$, per semplicità e per tenere un'uguale simbologia del dummy code si definisce effetto totale $\mu^{C=2\ A=2}$ e effetto diretto $\tau^{C=2\ A=2}$. La simbologia $\tau^{C=2\ A=2}$ è definita effetto diretto da Croon (2009). E' importante sottolineare che questa semplificazione non influisce sul risultato , utilizzato in seguito, che l'effetto sia maggiore od uguale ad 1: i numeri minori di uno se elevati al quadrato rimangono minori di uno, se maggiori rimangono maggiori.
Si ricava β e α in funzione dei parametri moltiplicativi $\mu^{C=2\ A=1}$, $\mu^{C=2}$ e $\mu^{C=2\ A=2}$

$$\beta = \frac{P(C=2|A=1)}{P(C=2|A=2)} = \frac{\mu^{C=2\ A=1}}{\mu^{C=2\ A=2}}\frac{\eta^{C|A=1}}{\eta^{C|A=2}}=(\mu^{C=2\ A=1})^2\left[(\mu^{C=2\ A=1})^2\frac{(\mu^{C=2}\mu^{C=2\ A=2})^2+1}{(\mu^{C=2}\mu^{C=2\ A=1})^2+1}\right]$$
(1.3)

$$\alpha = \frac{P(C=1|A=2)}{P(C=2|A=2)} = \frac{\mu^{C=1}\mu^{C=1\ A=2}}{\mu^{C=2}\mu^{C=2\ A=2}}=(\mu^{C=1\ A=2})^2(\mu^{C=2})^{-2}$$
(1.4)

dove per definizione, essendo costanti di normalizzazione, le η sono

$$\eta^{C|A=1} = \frac{\mu^{C=2}\mu^{C=2\ A=1}}{(\mu^{C=2}\mu^{C=2\ A=1})^2+1} \quad \text{e} \quad \eta^{C|A=2} = \frac{\mu^{C=2}\mu^{C=2\ A=2}}{(\mu^{C=2}\mu^{C=2\ A=2})^2+1} \quad => \quad \frac{\eta^{C|A=1}}{\eta^{C|A=2}}$$
$$=(\mu^{C=2\ A=1})^2\frac{(\mu^{C=2}\mu^{C=2\ A=2})^2+1}{(\mu^{C=2}\mu^{C=2\ A=1})^2+1}$$

Dalla (1.4) si ricava con semplici passaggi:

$$\mu^{C=2} = \frac{\mu^{C=1\ A=2}}{\sqrt[2]{\alpha}}$$

che inserita nell'equazione (1.3) risulta :

$$\beta = (\mu^{C=2\ A=1})^4 \frac{\left(\frac{1}{\sqrt[2]{\alpha}}\right)^2 + 1}{\left(\frac{(\mu^{C=2\ A=1})^2}{\sqrt[2]{\alpha}}\right)^2 + 1} = (\mu^{C=2\ A=1})^4 \frac{1+\alpha}{(\mu^{C=2\ A=1})^4 + \alpha}$$

Da questa relazione si ottiene la variazione $1/(\mu^{C=2\ A=1})^2$:

$$\beta[(\mu^{C=2\ A=1})^4 + \alpha] = (\mu^{C=2\ A=1})^4 \Rightarrow \frac{1}{(\mu^{C=2\ A=1})^2} = \sqrt{\frac{1+\alpha-\beta}{\alpha\beta}}$$

A.2 Effetto totale: calcolo diretto dai parametri stimati dal modello

Nella sezione precedente si è calcolato l'effetto totale come funzione delle probabilità condizionate C|A. Nei softwares queste condizionate non vengono calcolate quando si stima un modello causale a tre variabili. In questa sezione si sviluppa una formula per calcolare l'effetto totale direttamente dai parametri τ, sempre presenti nei risultati di stima. Ci si limita, per semplicità, a calcolare α e β come funzioni dei parametri moltiplicativi.Si calcola α :

$$\alpha = \frac{P(C=1|A=2)}{P(C=2|A=2)} \qquad (1.5)$$

Si calcola P(C=1|A=2) partendo dall'identità:

$$P(C=1|A=2) \equiv \frac{P(C=1,A=2,B=1)+P(C=1,A=2,B=2)}{P(A)} \qquad (1.6)$$

Sostituendo nella (1.6) le probabilità in forma moltiplicativa risulta P(C|A) in funzione dei parametri moltiplicativi τ:

$$P(C=1|A=2) = \frac{\eta^{C|A=2B=1}\eta^{B|A=2}\eta^{A=2}\,\tau^{A=2}\tau^{B=1}\tau^{C=1}\tau^{A=2\ B=1}\tau^{C=1\ A=2}\tau^{C=1B=1} + \eta^{C|A=2B=2}\eta^{B|A=2}\eta^{A=2}\,\tau^{A=2}\tau^{B=2}\tau^{C=1}\tau^{A=2\ B=2}\tau^{C=1\ A=2}\tau^{C=1B=2}}{\eta^{A=2}\,\tau^{A=2}}$$

Semplificando e raccogliendo si ottiene

$\eta^{B|A=2}\tau^{C=1}\tau^{C=1\ A=2}\left(\eta^{C|A=2B=1}\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=1B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{C=1B=2}\tau^{A=2\ B=2}\right)$
(1.7)

Applicando lo stesso procedimento a P(C=2|A=2), si ha P(C=2|A=2) in funzione dei parametri moltiplicativi τ:

$\eta^{B|A=2}\tau^{C=2}\tau^{C=2\ A=2}\left(\eta^{C|A=2\ B=1}\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=2\ B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\ B=2}\tau^{C=2\ B=2}\right)$
(1.8)

Sostituendo nella (1.6) le uguaglianze (1.7) e (1.8) risulta α in funzione dei parametri moltiplicativi τ:

$\alpha=\frac{\eta^{B|A=2}\tau^{C=1}\tau^{C=1\ A=2}\left(\eta^{C|A=2B=1}\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=1B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\ B=2}\tau^{C=1B=2}\right)}{\eta^{B|A=2}\tau^{C=2}\tau^{C=2\ A=2}\left(\eta^{C|A=2\ B=1}\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=2\ B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\ B=2}\tau^{C=2\ B=2}\right)}$
(1.9)

La stesso metodo si usa per ottenere β in funzione dei parametri moltiplicativi τ:

$\beta=\frac{P(C=2|A=1)}{P(C=2|A=2)}$
(2.0)

La probabilità condizionata $P(C=2|A=2)$è già stata calcolata nella (1.8) e con uguali passaggi si determina $P(C=2|A=1)$ in funzione dei parametri moltiplicativi τ:

$\eta^{B|A=1}\tau^{C=2}\tau^{C=2\ A=1}\left(\eta^{C|A=1\ B=1}\tau^{B=1}\tau^{A=1\ B=1}\tau^{C=2\ B=1}+\eta^{C|A=1\ B=2}\tau^{B=2}\tau^{A=1\ B=2}\tau^{C=2\ B=2}\right)$
(2.1)

Sostituendo nella (2.0) le uguaglianze (2.1) e (1.8) risulta

$\beta=\frac{\eta^{B|A=1}\tau^{C=2}\tau^{C=2\ A=1}\left(\eta^{C|A=1\ B=1}\tau^{B=1}\tau^{A=1\ B=1}\tau^{C=2\ B=1}+\eta^{C|A=1\ B=2}\tau^{B=2}\tau^{A=1\ B=2}\tau^{C=2\ B=2}\right)}{\eta^{B|A=2}\tau^{C=2}\tau^{C=2\ A=2}\left(\eta^{C|A=2\ B=1}\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=2\ B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\ B=2}\tau^{C=2\ B=2}\right)}$
(2.2)

Come fatto in precedenza, si effettua sia l'analisi in dummy code sia in effect code. La (2.2) e la (1.9) possono essere semplificate utilizzando la parametrizzazione dummy code o effect code:

- Dummy code

Ricordando che nella parametrizzazione dummy code le classi 1 delle variabili A, B e C sono baselines, si ha che

$\tau^{C=1} = \tau^{C=1\ A=2} = \tau^{C=1\ A=1} = \tau^{C=2\ A=1}$=1.

Inserendo la parametrizzazione dummy code nella (1.9), essa diviene:

$$\alpha=\frac{\left(\eta^{C|A=2B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\ B=2}\right)}{\tau^{C=2}\tau^{C=2\ A=2}\left(\eta^{C|A=2B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\ B=2}\tau^{C=2\ B=2}\right)}$$
(1.9d)

Inserendo la parametrizzazione dummy code nella (2.2), essa diviene:

$$\beta=\frac{\eta^{B|A=1}\left(\eta^{C|A=1\ B=1}+\eta^{C|A=1\ B=2}\tau^{B=2}\tau^{C=2\ B=2}\right)}{\eta^{B|A=2}\tau^{C=2\ A=2}\left(\eta^{C|A=2\ B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\ B=2}\tau^{C=2\ B=2}\right)} \qquad (2.2d)$$

Le equazioni (1.9d) e (2.2d) rappresentano gli effetti α e β in funzione dei parametri moltiplicativi τ e delle costanti η, le quali sono funzioni dei parametri moltiplicativi τ come definito nel capitolo 1.

- Effect code

Ricordando che nella parametrizzazione effect code:

$$\mu^{C=2\ A=2} = \mu^{C=1\ A=1} = \frac{1}{\mu^{C=2\ A=1}} = \frac{1}{\mu^{C=1\ A=2}}$$
$$\mu^{\ C=2} = \frac{1}{\mu^{C=1}}$$

Inserendo la parametrizzazione effect code nella (1.9), essa diviene:

$$\alpha=(\tau^{C=1})^2(\tau^{C=1\ A=2})^2\left[\frac{\left(\eta^{C|A=2B=1}\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=1B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\ B=2}\tau^{C=1B=2}\right)}{\left(\eta^{C|A=2\ B=1}\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=2\ B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\ B=2}\tau^{C=2\ B=2}\right)}\right]$$
(2.3)

La parte all'interno della parentesi quadrata può essere semplificata sia al numeratore che al denominatore

1)semplificazione numeratore

$$\eta^{C|A=2B=1}\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=1B=1} + \eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\ B=2}\tau^{C=1B=2} =$$

$$\eta^{C|A=2B=1}\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=1B=1}+\frac{\eta^{C|A=2B=2}}{\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=1\ B=1}}=$$
$$\frac{\eta^{C|A=2B=1}\left(\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=1\ B=1}\right)^2+\eta^{C|A=2B=2}}{\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=1\ B=1}}$$

2)semplificazione denominatore

Lo stesso metodo utilizzato per il numeratore viene applicato al denominatore che risulta:

$$\frac{\eta^{C|A=2B=1}(\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=2\ B=1})^2+\eta^{C|A=2B=2}}{\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=2\ B=1}}$$

Sostituendo le 2 semplificazioni nella (2.3), risulta:

α=
$$(\tau^{C=1})^2(\tau^{C=1\ A=2})^2\left[\frac{\eta^{C|A=2B=1}\left(\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=1\ B=1}\right)^2+\eta^{C|A=2B=2}}{\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=1\ B=1}}\right]\left[\frac{\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=2\ B=1}}{\eta^{C|A=2B=1}\left(\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=2\ B=1}\right)^2+\eta^{C|A=2B=2}}\right]$$
$$=(\tau^{C=1})^2(\tau^{C=1\ A=2})^2(\tau^{C=2\ A=1})^2\left[\frac{\eta^{C|A=2B=1}\left(\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=1\ B=1}\right)^2+\eta^{C|A=2B=2}}{\eta^{C|A=2B=1}\left(\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=2\ B=1}\right)^2+\eta^{C|A=2B=2}}\right] \qquad (1.9e)$$

Si passa alla semplificazione della funzione β (2.2)

$$\beta=\frac{\eta^{B|A=1}}{\eta^{B|A=2}}(\tau^{C=2\ A=1})^2\left[\frac{\left(\eta^{C|A=1\ B=1}\tau^{B=1}\tau^{A=1\ B=1}\tau^{C=2\ B=1}+\eta^{C|A=1\ B=2}\tau^{B=2}\tau^{A=1\ B=2}\tau^{C=2\ B=2}\right)}{\left(\eta^{C|A=2\ B=1}\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=2\ B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\ B=2}\tau^{C=2\ B=2}\right)}\right]$$
(2.4)

La parte all'interno della parentesi quadrata può essere semplificata sia al numeratore che al denominatore:

1)semplificazione numeratore

$$\eta^{C|A=1\ B=1}\tau^{B=1}\tau^{A=1\ B=1}\tau^{C=2\ B=1}+\eta^{C|A=1\ B=2}\tau^{B=2}\tau^{A=1\ B=2}\tau^{C=2\ B=2}=$$
$$\eta^{C|A=1\ B=1}\tau^{B=1}\tau^{A=1\ B=1}\tau^{C=2\ B=1}+$$
$$\frac{\eta^{C|A=1\ B=2}}{\tau^{B=1}\tau^{A=1\ B=1}\tau^{C=2\ B=1}}=\frac{\eta^{C|A=1\ B=1}\left(\tau^{B=1}\tau^{A=1\ B=1}\tau^{C=2\ B=1}\right)^2+\eta^{C|A=1\ B=2}}{\tau^{B=1}\tau^{A=1\ B=1}\tau^{C=2\ B=1}}$$

2)semplificazione denominatore
Il denominatore è lo stesso di α e perciò si utilizza la scomposizione appena svolta

Dalle 2 semplificazioni sostituite nella 2.4, si ottiene:

β=
$$\frac{\eta^{B|A=1}}{\eta^{B|A=2}}(\tau^{C=2\ A=1})^2\left[\frac{\eta^{C|A=1\ B=1}\left(\tau^{B=1}\tau^{A=1\ B=1}\tau^{C=2\ B=1}\right)^2+\eta^{C|A=1\ B=2}}{\tau^{B=1}\tau^{A=1\ B=1}\tau^{C=2\ B=1}}\right]\left[\frac{\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=2\ B=1}}{\eta^{C|A=2B=1}\left(\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=2\ B=1}\right)^2+\eta^{C|A=2B=2}}\right]$$

$$=\frac{\eta^{B|A=1}}{\eta^{B|A=2}}(\tau^{C=2\ A=1})^2(\tau^{A=2\ B=1})^2\left[\frac{\eta^{C|A=1\ B=1}\left(\tau^{B=1}\tau^{A=1\ B=1}\tau^{C=2\ B=1}\right)^2+\eta^{C|A=1\ B=2}}{\eta^{C|A=2B=1}\left(\tau^{B=1}\tau^{A=2\ B=1}\tau^{C=2\ B=1}\right)^2+\eta^{C|A=2B=2}}\right] \qquad (2.2e)$$

A.3 Scomposizione dell'effetto totale in effetto diretto e indiretto

La scomposizione dell'effetto totale in effetto diretto e indiretto totale (e questo a sua volta in effetto cella ed indiretto) è semplice solo nella formulazione dummy code, grazie al minor numero di parametri. Nel paper di Rao, Li e Roth (2008) l'effetto indiretto è definito come la differenza tra l'effetto totale e quello diretto ed, essendo utilizzata la formulazione log-lineare per gli effetti, è usata la differenza. In questa tesi si utilizzano sia i parametri log-lineari che quelli moltiplicativi, ma le formule degli effetti sono calcolate solo per i parametri moltiplicativi, in quanto la trasformazione per i parametri log-lineari è semplicemente una funzione logaritmica.

A.3.1 Effetto indiretto totale

Partendo da β (1.2) come funzione dell'effetto totale:

$$\beta=\frac{\eta^{C|A=1}}{\eta^{C|A=2}\mu^{C=2\ A=2}}$$

si ricava $\mu^{C=2\ A=2}$ in funzione di β:

$$\mu^{C=2\ A=2}=\frac{\eta^{C|A=1}}{\eta^{C|A=2}\beta} \qquad (2.5)$$

Per eliminare dalla (2.5) il rapporto $\eta^{C|A=1}/\eta^{C|A=2}$, valore che non è calcolabile dai parametri moltiplicativi stimati nel modello, si calcola il seguente rapporto:

$$\frac{P(C=1|A=1)}{P(C=1|A=2)}=\frac{\eta^{C|A=1}}{\eta^{C|A=2}}=\frac{\eta^{C|A=1B=1}+\eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=2B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\ B=2}}\frac{\eta^{B|A=1}}{\eta^{B|A=2}} \qquad (2.6)$$

Sostituendo nella (2.5) la (2.6) risulta:

$$\mu^{C=2\ A=2}=effetto_{totale}=\frac{1}{\beta}\frac{\eta^{C|A=1B=1}+\eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=2B=1}+\eta^{C|A=2\ B=2}\tau^{B=2}\tau^{A=2\ B=2}}\frac{\eta^{B|A=1}}{\eta^{B|A=2}} \qquad (2.7)$$

Sostituendo nella (2.7) il valore di β calcolato nella (2.2d) risulta:

$$effetto_{totale} = \left[\frac{\eta^{B|A=1}\left(\eta^{C|A=1\,B=1}+\eta^{C|A=1\,B=2}\tau^{B=2}\tau^{C=2\,B=2}\right)}{\eta^{B|A=2}\tau^{C=2\,A=2}\left(\eta^{C|A=2\,B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\,B=2}\tau^{C=2\,B=2}\right)}\right]^{-1}\left[\frac{\eta^{C|A=1B=1}+\eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=2B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\,B=2}}\,\frac{\eta^{B|A=1}}{\eta^{B|A=2}}\right]$$

formulazione che può essere semplificata e riordinata per permettere una facile individuazione delle due parti:

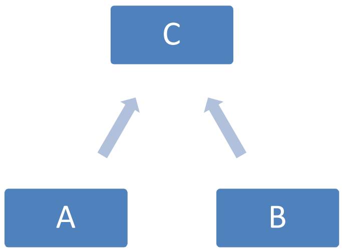
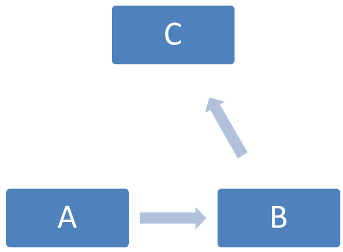
$$effetto_{totale} = \tau^{C=2\,A=2}\left\{\left[\frac{\eta^{C|A=1B=1}+\eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=2B=1}+\eta^{C|A=2\,B=2}\tau^{B=2}\tau^{A=2\,B=2}}\right]\left[\frac{\left(\eta^{C|A=1\,B=1}+\eta^{C|A=1\,B=2}\tau^{B=2}\tau^{C=2\,B=2}\right)}{\left(\eta^{C|A=2\,B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\,B=2}\tau^{C=2\,B=2}\right)}\right]^{-1}\right\} \quad (2.8)$$

Dove
$\tau^{C=2\,A=2}$ è l'effetto diretto

$\left\{\left[\frac{\eta^{C|A=1B=1}+\eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=2B=1}+\eta^{C|A=2\,B=2}\tau^{B=2}\tau^{A=2\,B=2}}\right]\left[\frac{\left(\eta^{C|A=1\,B=1}+\eta^{C|A=1\,B=2}\tau^{B=2}\tau^{C=2\,B=2}\right)}{\left(\eta^{C|A=2\,B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\,B=2}\tau^{C=2\,B=2}\right)}\right]^{-1}\right\}$ è l'effetto indiretto totale

A.3.2 Effetto cella

L'effetto cella risulta dal processo generatore dei dati ed è una particolare interazione tra gli effetti diretti che influiscono su di una medesima variabile. Con una semplice rappresentazione grafica si può spiegare la sua presenza:

Modello con solo effetto cella	Modello senza effetto cella
	

Nel grafico di sinistra ci sono 2 variabili, A e B, che influiscono direttamente su C, quindi le η che costruiscono la condizionata C|AB sono 4:

$\eta^{C|A=1\,B=1}$ garantisce che P(C=1|A=1,B=1)+P(C=2|A=1,B=1)=1

$\eta^{C|A=2\,B=1}$ garantisce che P(C=1|A=2,B=1)+P(C=2|A=2,B=1)=1

$\eta^{C|A=1\,B=2}$ garantisce che P(C=1|A=1,B=2)+P(C=2|A=1,B=2)=1

$\eta^{C|A=2\,B=2}$ garantisce che P(C=1|A=2,B=2)+P(C=2|A=2,B=2)=1

Nel grafico di destra c'è solo la variabile B che influisce su C, perciò le η che costruiscono la condizionata C|B sono 2:

$\eta^{C|B=1}$ garantisce che P(C=1|B=1)+P(C=2|B=1)=1

$\eta^{C|B=2}$ garantisce che P(C=1|B=2)+P(C=2|B=2)=1

Tutte le volte che più di una variabile influisce su C, c'è l'effetto cella. L'effetto cella si calcola dal modello di sinistra dove $\tau^{A=2\ B=2}$=1,

infatti nel modello di sinistra

$$effetto_{totale} = effetto\ diretto + effetto\ cella$$

Inserendo nella (2.8) la restrizione del modello di sinistra si ottiene:

$$effetto_{totale} = \tau^{C=2\ A=2}\left\{\left[\frac{\eta^{C|A=1B=1} + \eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=2B=1} + \eta^{C|A=2\ B=2}\tau^{B=2}}\right]\left[\frac{\left(\eta^{C|A=1\ B=1} + \eta^{C|A=1\ B=2}\tau^{B=2}\tau^{C=2\ B=2}\right)}{\left(\eta^{C|A=2\ B=1} + \eta^{C|A=2B=2}\tau^{B=2}\tau^{C=2\ B=2}\right)}\right]^{-1}\right\}$$

dove:

$\tau^{C=2\ A=2} = effetto\ diretto$

$\left\{\left[\frac{\eta^{C|A=1B=1}+\eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=2B=1}+\eta^{C|A=2\ B=2}\tau^{B=2}}\right]\left[\frac{\left(\eta^{C|A=1\ B=1}+\eta^{C|A=1\ B=2}\tau^{B=2}\tau^{C=2\ B=2}\right)}{\left(\eta^{C|A=2\ B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{C=2\ B=2}\right)}\right]^{-1}\right\}$= effetto cella

L'effetto cella è una particolare interazione tra i 2 effetti diretti $\tau^{C=2\ B=2}$ e $\tau^{C=2\ A=2}$. La sua trasformazione logaritmica ha segno positivo se $\tau^{C=2\ A=2} < 1$ e negativo se $\tau^{C=2\ A=2} > 1$. Rappresenta la presenza di A che influisce sull'effetto indiretto o la presenza B che influisce sull'effetto diretto. Il segno dell'effetto cella è determinato solo da $\tau^{C=2\ A=2}$, mentre la sua intensità congiuntamente da $\tau^{C=2\ B=2}$ e da $\tau^{C=2\ A=2}$. Naturalmente in termini di parametri moltiplicativi con positivo si intende maggiore di 1, con negativo minore di 1. La dimostrazione è la seguente.

Sostituendo le costanti η nella formula dell'effetto cella risulta:

$$effetto_{cella} = \frac{\left(1+\tau^{B=2}\right)+\tau^{C=2}\left(\tau^{C=2\ B=2}+\tau^{B=2}\right)}{\left(1+\tau^{B=2}\tau^{C=2\ B=2}\right)+\tau^{C=2}\tau^{C=2\ B=2}\left(1+\tau^{B=2}\right)}\ \frac{\left(1+\tau^{B=2}\tau^{C=2\ B=2}\right)+\tau^{C=2}\tau^{C=2\ A=2}\tau^{C=2\ B=2}\left(1+\tau^{B=2}\right)}{\left(1+\tau^{B=2}\right)+\tau^{C=2}\tau^{C=2\ A=2}\left(\tau^{C=2\ B=2}+\tau^{B=2}\right)}$$

Si studia quando l'effetto cella è maggiore di 1:

$$\frac{\left(1+\tau^{B=2}\right)+\tau^{C=2}\left(\tau^{C=2\ B=2}+\tau^{B=2}\right)}{\left(1+\tau^{B=2}\tau^{C=2\ B=2}\right)+\tau^{C=2}\tau^{C=2\ B=2}\left(1+\tau^{B=2}\right)}\ \frac{\left(1+\tau^{B=2}\tau^{C=2\ B=2}\right)+\tau^{C=2}\tau^{C=2\ A=2}\tau^{C=2\ B=2}\left(1+\tau^{B=2}\right)}{\left(1+\tau^{B=2}\right)+\tau^{C=2}\tau^{C=2\ A=2}\left(\tau^{C=2\ B=2}+\tau^{B=2}\right)} > 1$$

Il denominatore è sempre diverso da 0 e sempre positivo perché somma di valori positivi:

$$[(1+\tau^{B=2})+\tau^{C=2}(\tau^{C=2\,B=2}+\tau^{B=2})][(1+\tau^{B=2}\tau^{C=2\,B=2})+\tau^{C=2}\tau^{C=2\,A=2}\tau^{C=2\,B=2}(1+\tau^{B=2})]>$$
$$[(1+\tau^{B=2}\tau^{C=2\,B=2})+\tau^{C=2}\tau^{C=2\,B=2}(1+\tau^{B=2})][(1+\tau^{B=2})+\tau^{C=2}\tau^{C=2\,A=2}(\tau^{C=2\,B=2}+\tau^{B=2})]$$

=>

$$(1+\tau^{B=2})(1+\tau^{B=2}\tau^{C=2\,B=2})+(1+\tau^{B=2})\tau^{C=2}\tau^{C=2\,A=2}\tau^{C=2\,B=2}(1+\tau^{B=2})+\tau^{C=2}(\tau^{C=2\,B=2}+\tau^{B=2})$$
$$(1+\tau^{B=2}\tau^{C=2\,B=2})+\tau^{C=2}(\tau^{C=2\,B=2}+\tau^{B=2})\tau^{C=2}\tau^{C=2\,A=2}\tau^{C=2\,B=2}(1+\tau^{B=2})>$$
$$(1+\tau^{B=2}\tau^{C=2\,B=2})(1+\tau^{B=2})+\tau^{C=2}\tau^{C=2\,B=2}(1+\tau^{B=2})(1+\tau^{B=2})+$$
$$(1+\tau^{B=2}\tau^{C=2\,B=2})\tau^{C=2}\tau^{C=2\,A=2}(\tau^{C=2\,B=2}+\tau^{B=2})+\tau^{C=2}\tau^{C=2\,B=2}\left(1+\tau^{B=2})\tau^{C=2}\tau^{C=2\,A=2}(\tau^{C=2\,B=2}+\tau^{B=2})\right)$$

semplificando risulta:

$$(1+\tau^{B=2})\tau^{C=2}\tau^{C=2\,A=2}\tau^{C=2\,B=2}(1+\tau^{B=2})+\tau^{C=2}(\tau^{C=2\,B=2}+\tau^{B=2})(1+\tau^{B=2}\tau^{C=2\,B=2})>$$
$$\tau^{C=2}\tau^{C=2\,B=2}(1+\tau^{B=2})(1+\tau^{B=2})+(1+\tau^{B=2}\tau^{C=2\,B=2})\tau^{C=2}\tau^{C=2\,A=2}(\tau^{C=2\,B=2}+\tau^{B=2})$$

semplificando ulteriormente:

$$(1+\tau^{B=2})^2\tau^{C=2\,A=2}\tau^{C=2\,B=2}+(1+\tau^{B=2}\tau^{C=2\,B=2})(\tau^{C=2\,B=2}+\tau^{B=2})$$
$$>(1+\tau^{B=2})^2\tau^{C=2\,B=2}+(1+\tau^{B=2}\tau^{C=2\,B=2})\tau^{C=2\,A=2}(\tau^{C=2\,B=2}+\tau^{B=2})$$

e raccogliendo

$$(1+\tau^{B=2})^2\tau^{C=2\,B=2}(\tau^{C=2\,A=2}-1)-(1+\tau^{B=2}\tau^{C=2\,B=2})(\tau^{C=2\,B=2}+\tau^{B=2})(\tau^{C=2\,A=2}-1)>0$$
=>
$$(\tau^{C=2\,A=2}-1)[(1+\tau^{B=2})^2\tau^{C=2\,B=2}-(1+\tau^{B=2}\tau^{C=2\,B=2})(\tau^{C=2\,B=2}+\tau^{B=2})]>0$$
=>
$$(\tau^{C=2\,A=2}-1)\left[\tau^{C=2\,B=2}+2\tau^{C=2\,B=2}\tau^{B=2}+{\tau^{B=2}}^2\tau^{C=2\,B=2}-\tau^{C=2\,B=2}-\tau^{B=2}-\tau^{B=2}{\tau^{C=2\,B=2}}^2\right.$$
$$\left.-{\tau^{B=2}}^2\tau^{C=2\,B=2}\right]>0$$
=>
$$(\tau^{C=2\,A=2}-1)\left[2\tau^{C=2\,B=2}\tau^{B=2}-\tau^{B=2}-\tau^{B=2}{\tau^{C=2\,B=2}}^2\right]>0$$
=>
$$\tau^{B=2}(\tau^{C=2\,A=2}-1)\left[2\tau^{C=2\,B=2}-1-{\tau^{C=2\,B=2}}^2\right]>0$$
=>
$$-\tau^{B=2}(\tau^{C=2\,A=2}-1)[\tau^{C=2\,B=2}-1]^2>0$$

Affinché $-\tau^{B=2}(\tau^{C=2\,A=2}-1)[\tau^{C=2\,B=2}-1]^2$ sia maggiore di 0 è necessario che $\tau^{C=2\,A=2}<1$ che in termini di parametri log-lineari significa che l'effetto diretto è negativo. Nella figura 1 si vede come influiscono i 2 parametri sull'effetto cella. I parametri sono quelli usati nel capitolo 1 "modello completo" tabella 22, ad eccezione di $\tau^{C=2\,A=2}$ e di $\tau^{C=2\,B=2}$ che variano. L'effetto cella del modello completo del capitolo 1 si trova nella curva rossa ($\tau^{C=2\,B=2}=0.5$) con $\tau^{C=2\,A=2}$=0.8. Si nota dal grafico che un aumento di $\tau^{C=2\,A=2}$ porta ad una diminuzione dell'effetto cella, la cui intensità dipende da $\tau^{C=2\,B=2}$.

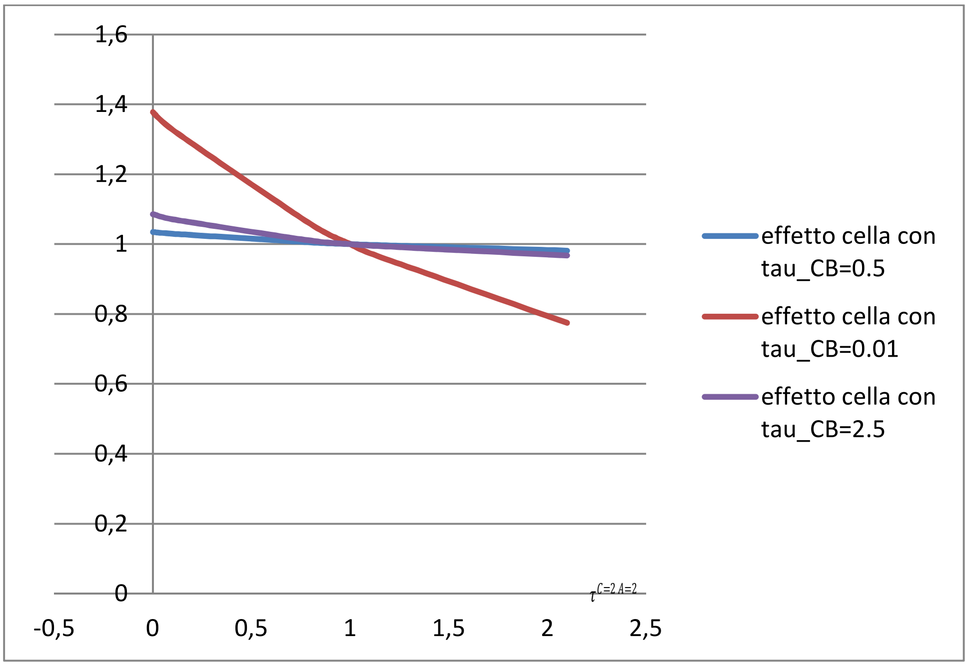


Figura 1: effetto cella in funzione dell'effetto diretto di A su C

La relazione tra il parametro dell'effetto diretto $\tau^{C=2\;A=2}$ e l'effetto cella si generalizza studiando il segno della derivata prima dell'effetto cella rispetto a $\tau^{C=2\;A=2}$.

$$effetto_{cella} = \frac{(1+\tau^{B=2}) + \tau^{C=2}(\tau^{C=2\;B=2}+\tau^{B=2})}{(1+\tau^{B=2}\tau^{C=2\;B=2}) + \tau^{C=2}\tau^{C=2\;B=2}(1+\tau^{B=2})} \; \frac{(1+\tau^{B=2}\tau^{C=2\;B=2}) + \tau^{C=2}\tau^{C=2\;A=2}\tau^{C=2\;B=2}(1+\tau^{B=2})}{(1+\tau^{B=2}) + \tau^{C=2}\tau^{C=2\;A=2}(\tau^{C=2\;B=2}+\tau^{B=2})}$$
$$= \mathrm{H}\,(\tau^{C=2\;B=2}, \tau^{B=2}, \tau^{C=2})\; \mathrm{G}(\tau^{C=2\;B=2}, \tau^{B=2}, \tau^{C=2}, \tau^{C=2\;A=2})$$

=>

$$\frac{\partial\, effetto_{cella}}{\partial \tau^{C=2\;A=2}} = \mathrm{H}\,(\tau^{C=2\;B=2}, \tau^{B=2}, \tau^{C=2})\, \frac{\partial \mathrm{G}(\tau^{C=2\;B=2}, \tau^{B=2}, \tau^{C=2}, \tau^{C=2\;A=2})}{\partial \tau^{C=2\;A=2}}$$

La funzione H $(\tau^{C=2\;B=2}, \tau^{B=2}, \tau^{C=2})$ è sempre positiva e quindi il segno della derivata dipende dalla derivata $(\partial \mathrm{G}(\tau^{C=2\;B=2}, \tau^{B=2}, \tau^{C=2}, \tau^{C=2\;A=2})/\partial \tau^{C=2\;A=2})$ della funzione G rispetto a $\tau^{C=2\;A=2}$:

$$\frac{\partial \mathrm{G}(\tau^{C=2\;B=2}, \tau^{B=2}, \tau^{C=2}, \tau^{C=2\;A=2})}{\partial \tau^{C=2\;A=2}} = \frac{\tau^{C=2}\tau^{C=2\;B=2}(1+\tau^{B=2})[(1+\tau^{B=2}) + \tau^{C=2}\tau^{C=2\;A=2}(\tau^{C=2\;B=2}+\tau^{B=2})]}{[(1+\tau^{B=2}) + \tau^{C=2}\tau^{C=2\;A=2}(\tau^{C=2\;B=2}+\tau^{B=2})]^2} -$$
$$+ \frac{\tau^{C=2}(\tau^{C=2\;B=2}+\tau^{B=2})[(1+\tau^{B=2}\tau^{C=2\;B=2}) + \tau^{C=2}\tau^{C=2\;A=2}\tau^{C=2\;B=2}(1+\tau^{B=2})]}{[(1+\tau^{B=2}) + \tau^{C=2}\tau^{C=2\;A=2}(\tau^{C=2\;B=2}+\tau^{B=2})]^2}$$
$$= \frac{\tau^{C=2}\tau^{C=2\;B=2}(1+\tau^{B=2})^2 - \tau^{C=2}(\tau^{C=2\;B=2}+\tau^{B=2})[(1+\tau^{B=2}\tau^{C=2\;B=2})]}{[(1+\tau^{B=2}) + \tau^{C=2}\tau^{C=2\;A=2}(\tau^{C=2\;B=2}+\tau^{B=2})]^2}$$

Il segno della derivata della funzione G è determinato dal numeratore perché il denominatore è sempre maggiore di 0 essendo un quadrato

$$\tau^{C=2}\tau^{C=2\ B=2}(1+\tau^{B=2})^2-\tau^{C=2}(\tau^{C=2\ B=2}+\tau^{B=2})[(1+\tau^{B=2}\tau^{C=2\ B=2})]>0$$

Semplificando il parametro $\tau^{C=2}$, che è sempre positivo per definizione, e sviluppando il quadrato, si ottiene:

$\tau^{C=2\ B=2}\left(1+2\tau^{B=2}+{\tau^{B=2}}^2\right)-\ (\tau^{C=2\ B=2}+\tau^{B=2}+\tau^{B=2}(\tau^{C=2\ B=2})^2+$
$(\tau^{B=2})^2\tau^{C=2\ B=2})>0$

=>
$-\tau^{B=2}(1-\tau^{C=2\ B=2})^2>0$

La disequazione $-\tau^{B=2}(1-\tau^{C=2\ B=2})^2>0$ non è mai verificata mentre è sempre verificata la disequazione $-\tau^{B=2}(1-\tau^{C=2\ B=2})^2<0$, quindi la derivata prima dell'effetto cella rispetto a $\tau^{C=2\ A=2}$ è sempre negativa. Conseguentemente un aumento dell'effetto diretto $\tau^{C=2\ A=2}$porta ad una diminuzione dell'effetto cella.

A.3.3 Effetto indiretto

Ricordando che l'effetto indiretto totale è il prodotto tra l'effetto cella e quello indiretto, risulta:

Effetto_indiretto=effetto_ind_tot/effetto_cella

Da questa formula si può studiare il segno dell'effetto indiretto. Si studia quando l'effetto indiretto è minore di uno:

effetto_indiretto= effetto_ind_tot/effetto_cella<1
=> effetto_ind_tot<effetto_cella

Partendo da questa disuguaglianza e sostituendo in essa le rispettive formule dell'effetto cella e di quello indiretto_totale, risulta:

$$\left\{\left[\frac{\eta^{C|A=1B=1}+\eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=2B=1}+\eta^{C|A=2\ B=2}\tau^{B=2}\tau^{A=2\ B=2}}\right]\left[\frac{\left(\eta^{C|A=1\ B=1}+\eta^{C|A=1\ B=2}\tau^{B=2}\tau^{C=2\ B=2}\right)}{\left(\eta^{C|A=2\ B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\ B=2}\tau^{C=2\ B=2}\right)}\right]^{-1}\right\}<$$
$$\left\{\left[\frac{\eta^{C|A=1B=1}+\eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=2B=1}+\eta^{C|A=2\ B=2}\tau^{B=2}}\right]\left[\frac{\left(\eta^{C|A=1\ B=1}+\eta^{C|A=1\ B=2}\tau^{B=2}\tau^{C=2\ B=2}\right)}{\left(\eta^{C|A=2\ B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{C=2\ B=2}\right)}\right]^{-1}\right\}$$

Semplificando risulta:

$$\frac{\eta^{C|A=2B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{C=2\,B=2}\tau^{A=2\,B=2}}{\eta^{C|A=2B=1}+\eta^{C|A=2\,B=2}\tau^{B=2}\tau^{A=2\,B=2}} < \frac{\eta^{C|A=2B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{C=2\,B=2}}{\eta^{C|A=2B=1}+\eta^{C|A=2\,B=2}\tau^{B=2}}$$

Facendo comun denominatore e semplificando essendo il denominatore sempre maggiore di 0, risulta

$$\left(\eta^{C|A=2B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{C=2\,B=2}\tau^{A=2\,B=2}\right)\left(\eta^{C|A=2B=1}+\eta^{C|A=2\,B=2}\tau^{B=2}\right) < \left(\eta^{C|A=2B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{C=2\,B=2}\right)\left(\eta^{C|A=2B=1}+\eta^{C|A=2\,B=2}\tau^{B=2}\tau^{A=2\,B=2}\right)$$

Semplificando e raccogliendo le η risulta

$\tau^{B=2}\left(\eta^{C|A=2B=1}\eta^{C|A=2B=1}\right)(-\tau^{A=2\,B=2}-\tau^{C=2\,B=2}+1+\tau^{C=2\,B=2}\tau^{A=2\,B=2})<0$

Il segno è determinato esclusivamente da $\left(-\tau^{A=2\,B=2}-\tau^{C=2\,B=2}+1+\tau^{C=2\,B=2}\tau^{A=2\,B=2}\right)$,

$\left(-\tau^{A=2\,B=2}-\tau^{C=2\,B=2}+1+\tau^{C=2\,B=2}\tau^{A=2\,B=2}\right) = -\tau^{A=2\,B=2}(+1-\tau^{C=2\,B=2}) + (+1-\tau^{C=2\,B=2})=$

$=(+1-\tau^{C=2\,B=2})(+1-\tau^{A=2\,B=2})< 0$

che è verificata per $\tau^{C=2\,B=2} > 0$ e $\tau^{A=2\,B=2} < 0$ o per $\tau^{C=2\,B=2} < 0$ e $\tau^{A=2\,B=2} > 0$.

L'effetto indiretto si comporta come l'effetto indiretto della letteratura, cioè:

1) se il parametro A→B è positivo e B→C è negativo, l'effetto indiretto è negativo
2) se il parametro A→B è negativo e B→C è positivo, l'effetto indiretto è negativo
3) se il parametro A→B è positivo e B→C è positivo, l'effetto indiretto è positivo
4) se il parametro A→B è negativo e B→C è negativo, l'effetto indiretto è positivo

A.4 Ad ogni parametro il suo effetto

Come sintetizzato nella seguente tabella, la presenza/assenza di un parametro dal modello causa la presenza/l'assenza dell'effetto. Di seguito dimostro quali parametri causino l'effetto, mostrando che la loro assenza rende nullo l'effetto da loro causato.

Parametro/i	effetto
$\tau^{C=2\,A=2}$	diretto
$\tau^{C=2\,B=2}$ e $\tau^{A=2\,B=2}$	indiretto
1)$\tau^{C=2\,B=2}$ 2)$\tau^{C=2\,A=2}$	cella

- **Dimostrazione:effetto diretto**

Per controllare se $\tau^{C=2\,A=2}$ è il parametro di interesse dell'effetto diretto, si prova a inserirlo uguale ad 1 (cioè assenza d'effetto) nella formula (2.8):

$$effetto_{totale} = \left\{ \left[\frac{\eta^{C|A=1B=1} + \eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=2B=1} + \eta^{C|A=2\,B=2}\tau^{B=2}\tau^{A=2\,B=2}} \right] \left[\frac{\left(\eta^{C|A=1\,B=1} + \eta^{C|A=1\,B=2}\tau^{B=2}\tau^{C=2\,B=2}\right)}{\left(\eta^{C|A=2\,B=1} + \eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\,B=2}\tau^{C=2\,B=2}\right)} \right]^{-1} \right\} = effetto_{ind\_totale}$$

(2.9)

L'effetto totale, in assenza di $\tau^{C=2\,A=2}$, è composto solo dall'effetto_ind_tot, perciò l'effetto diretto sparisce.

- **Dimostrazione:effetto indiretto**

Per controllare se $\tau^{C=2\,B=2}$ è il parametro di interesse dell'effetto indiretto, si prova a inserirlo uguale ad 1 (cioè assenza d'effetto) nella formula (2.8):

$$effetto_{totale} = \tau^{C=2\,A=2} \left\{ \left[\frac{\eta^{C|A=1B=1} + \eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=2B=1} + \eta^{C|A=2\,B=2}\tau^{B=2}\tau^{A=2\,B=2}} \right] \left[\frac{\left(\eta^{C|A=1\,B=1} + \eta^{C|A=1\,B=2}\tau^{B=2}\right)}{\left(\eta^{C|A=2\,B=1} + \eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\,B=2}\right)} \right]^{-1} \right\}$$

(3.0)

Ricordando come si sono calcolate le costanti di normalizzazione η, risulta:

$$\eta^{C|A=1B=1} = \frac{1}{1+\tau^{C=2}}$$

$$\eta^{C|A=1B=2} = \frac{1}{1+\tau^{C=2}\,\tau^{C=2\,B=2}} = \frac{1}{1+\tau^{C=2}}$$

$$\eta^{C|A=2\,B=1} = \frac{1}{1+\tau^{C=2}\,\tau^{C=2\,A=2}}$$

$$\eta^{C|A=2\,B=2} = \frac{1}{1+\tau^{C=2}\,\tau^{C=2\,A=2}\tau^{C=2\,B=2}} = \frac{1}{1+\tau^{C=2}\,\tau^{C=2\,A=2}}$$

Essendo $\eta^{C|A=1B=1} = \eta^{C|A=1B=2}$ e $\eta^{C|A=2\,B=1} = \eta^{C|A=2\,B=2}$ si inseriscono queste uguaglianze nella (3.0) e risulta:

$$effetto_{totale} = \tau^{C=2\ A=2} \left\{ \left[\frac{\eta^{C|A=1B=1} + \eta^{C|A=1B=1}\tau^{B=2}}{\eta^{C|A=2B=2} + \eta^{C|A=2\ B=2}\tau^{B=2}\tau^{A=2\ B=2}} \right] \left[\frac{(\eta^{C|A=1\ B=1} + \eta^{C|A=1\ B=1}\tau^{B=2})}{(\eta^{C|A=2\ B=2} + \eta^{C|A=2B=2}\tau^{B=2}\tau^{A=2\ B=2})} \right]^{-1} \right\} = \tau^{C=2\ A=2}$$

L’effetto totale, in assenza di $\tau^{C=2\ B=2}$, è composto solo dall’effetto diretto, perciò l’effetto indiretto sparisce.

- **Dimostrazione: Effetto cella**

L’effetto cella dipende da due parametri: $\tau^{C=2\ A=2}$ e $\tau^{C=2\ B=2}$. Richiamando la formula dell’effetto cella:

$$effetto_{cella} = \left\{ \left[\frac{\eta^{C|A=1B=1} + \eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=2B=1} + \eta^{C|A=2\ B=2}\tau^{B=2}} \right] \left[\frac{(\eta^{C|A=1\ B=1} + \eta^{C|A=1\ B=2}\tau^{B=2}\tau^{C=2\ B=2})}{(\eta^{C|A=2\ B=1} + \eta^{C|A=2B=2}\tau^{B=2}\tau^{C=2\ B=2})} \right]^{-1} \right\}$$

(3.1)

- Per controllare se $\tau^{C=2\ A=2}$ è il parametro di interesse dell’effetto cella, si prova a inserirlo uguale ad 1 (cioè assenza d’effetto) nella formula (3.1). Il parametro $\tau^{C=2\ A=2}$ non si vede direttamente nella formula (3.1), ma è nei valori delle costanti di normalizzazione η:

$$\eta^{C|A=1B=1} = \frac{1}{1+\tau^{C=2}}$$

$$\eta^{C|A=1B=2} = \frac{1}{1+\tau^{C=2}\ \tau^{C=2\ B=2}}$$

$$\eta^{C|A=2\ B=1} = \frac{1}{1+\tau^{C=2}\ \tau^{C=2\ A=2}} = \frac{1}{1+\tau^{C=2}}$$

$$\eta^{C|A=2\ B=2} = \frac{1}{1+\tau^{C=2}\ \tau^{C=2\ A=2}\tau^{C=2\ B=2}} = \frac{1}{1+\tau^{C=2}\ \tau^{C=2\ B=2}}$$

Si ricavano le seguenti uguaglianze $\eta^{C|A=1B=1} = \eta^{C|A=2\ B=1}$ e $\eta^{C|A=1\ B=2} = \eta^{C|A=2\ B=2}$ e si inseriscono nella (3.1) ottenendo:

$$effetto_{cella} = \left\{ \left[\frac{\eta^{C|A=1B=1} + \eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=1B=1} + \eta^{C|A=1\ B=2}\tau^{B=2}} \right] \left[\frac{(\eta^{C|A=1\ B=1} + \eta^{C|A=1\ B=2}\tau^{B=2}\tau^{C=2\ B=2})}{(\eta^{C|A=1\ B=1} + \eta^{C|A=1B=2}\tau^{B=2}\tau^{C=2\ B=2})} \right]^{-1} \right\} = 1$$

L’effetto cella, in assenza di $\tau^{C=2\ A=2}$, essendo uguale ad 1,sparisce.

- Per controllare se $\tau^{C=2\ B=2}$ è il parametro di interesse dell’effetto cella, si prova a inserirlo uguale ad 1 (cioè assenza d’effetto) nella formula (3.1):

$$effetto_{cella} = \left\{ \left[\frac{\eta^{C|A=1B=1} + \eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=2B=1} + \eta^{C|A=2\ B=2}\tau^{B=2}} \right] \left[\frac{(\eta^{C|A=1\ B=1} + \eta^{C|A=1\ B=2}\tau^{B=2})}{(\eta^{C|A=2\ B=1} + \eta^{C|A=2B=2}\tau^{B=2})} \right]^{-1} \right\} = 1$$

A.5 Effetto totale nullo

Nel capitolo 5 del libro “Marginal Model”, Croon et al. (2009) affermano che in un modello generale può esserci, a causa dei segni opposti degli effetti diretti e indiretti, un effetto totale nullo che può essere inserito nel processo di stima (gli autori utilizzano il programma R) ponendo la condizione di indipendenza nella tabella di frequenza. Nel caso a 3 variabili, qui analizzato, si deve porre l’indipendenza tra A e C nella tabella di frequenza AC, che viene stimata dall’equazione:

$$\pi^{C|A} = \eta^{C|A} \mu^{C\,A} \mu^{C} \qquad (3.2)$$

La (3.2) è la formula che contiene l’effetto totale $\mu^{C\,A}$. Imporre che C e A siano indipendenti significa porre $\mu^{C\,A}$ uguale ad 1 che si ha solo quando β è uguale ad 1, sia nel caso effect code che nel caso dummy code. La dimostrazione è la seguente.

- Dummy code

Riprendendo la formula dell’effetto totale:

$$\mu^{A=2\ C=2} = \frac{\alpha+1-\beta}{\alpha\beta} \qquad (3.3)$$

e ponendo β uguale ad 1 nella (3.3), risulta

$$\mu^{A=2\ C=2} = \frac{\alpha+1-\beta}{\alpha\beta} = \frac{\alpha}{\alpha} = 1$$

- Effect code

Riprendendo la formula dell’effetto totale:

$$\mu^{A=2\ C=2} = \sqrt[4]{\frac{\alpha+1-\beta}{\alpha\beta}} \qquad (3.4)$$

e ponendo β uguale ad 1 nella (3.3), risulta

$$\mu^{A=2\ C=2} = \sqrt[4]{\frac{\alpha+1-\beta}{\alpha\beta}} = \sqrt[4]{\frac{\alpha}{\alpha}} = 1$$

A.6 Passaggio da effect code a dummy code

Il dummy code e l’effect code sono due parametrizzazioni utilizzate per convenzione per identificare un modello. I parametri stimati con un metodo possono essere ricondotti a quelli stimati con l'altro metodo. In questa sezione dell'Appendice si mostra il passaggio da parametri "effect code" a parametri "dummy code", prima con una dimostrazione matematica e poi utilizzando dei dati. La scelta è dovuta al fatto che sia le formule dell'effetto indiretto che quelle dell'effetto cella sono funzioni dei parametri in dummy code. Per distinguere i parametri effect code da quelli dummy code si utilizza la seguente notazione:

ϑ=parametri effect code ω= costante di normalizzazione effect code

τ=parametri dummy code η= costante di normalizzazione dummy code

I parametri calcolati sono quelli di un modello completo causale a 3 variabili dicotomiche dove A è esogena e B e C sono endogene, come nel capitolo 1 (1.4, 1.5)

- $\boldsymbol{parametro\ \tau^{A=2}}$

Per calcolare il parametro $\tau^{A=2}$ si parte dalla probabilità P(A=1)

$$P(A=1)=\eta^A\tau^{A=1} = \omega^A\vartheta^{A=1} \qquad (3.5)$$

ricordando che le costanti di normalizzazione sono:

$$\eta^A = \frac{1}{1+\tau^{A=2}}$$

$$\omega^A = \frac{1}{\vartheta^{A=1}+\vartheta^{A=2}} = \frac{\vartheta^{A=1}}{1+(\vartheta^{A=1})^2}$$

la (3.5) diviene

$$\frac{1}{1+\tau^{A=2}} = \frac{\vartheta^{A=1}}{1+(\vartheta^{A=1})^2}\ \vartheta^{A=1}$$

e sviluppando risulta

$$1+\tau^{A=2} = \frac{1+(\vartheta^{A=1})^2}{(\vartheta^{A=1})^2} = \frac{1}{(\vartheta^{A=1})^2} + 1$$

Il parametro $\tau^{A=2}$ è

$$\tau^{A=2} = \frac{1}{(\vartheta^{A=1})^2}$$

- $\boldsymbol{parametro\ \tau^{B=2}}$

Per calcolare il parametro $\tau^{B=2}$ si parte dalla probabilità P(B=1|A=1)

$$P(B=1|A=1)=\eta^{B|A=1}\tau^{B=1}\,\tau^{B=1\ A=1} = \omega^{B|A=1}\,\vartheta^{B=1}\vartheta^{B=1\ A=1} \qquad (3.6)$$

ricordando che le costanti di normalizzazione sono:

$$\eta^{B|A=1} = \frac{1}{1+\tau^{B=2}}$$

$$\omega^{B|A=1} = \frac{1}{\vartheta^{B=1}\vartheta^{B=1\ A=1} + \vartheta^{B=2}\vartheta^{B=2\ A=1}} = \frac{\vartheta^{B=1}\vartheta^{B=1\ A=1}}{1+(\vartheta^{B=1}\vartheta^{B=1\ A=1})^2}$$

la (3.6) diviene

$$\frac{1}{1+\tau^{B=2}} = \frac{\vartheta^{B=1}\vartheta^{B=1\ A=1}}{1+(\vartheta^{B=1}\vartheta^{B=1\ A=1})^2}\ \vartheta^{B=1}\vartheta^{B=1\ A=1}$$

e sviluppando risulta

$$1+\tau^{B=2} = \frac{1+(\vartheta^{B=1}\vartheta^{B=1\ A=1})^2}{(\vartheta^{B=1}\vartheta^{B=1\ A=1})^2} = 1 + \frac{1}{(\vartheta^{B=1}\vartheta^{B=1\ A=1})^2}$$

Il parametro $\tau^{B=2}$ è

$$\tau^{B=2} = \frac{1}{(\vartheta^{B=1}\vartheta^{B=1\ A=1})^2}$$

- ***parametro*** $\boldsymbol{\tau^{B=2\ A=2}}$

Per calcolare il parametro $\tau^{B=2\ A=2}$ si parte dalla probabilità P(B=1|A=2)

P(B=1|A=2)=$\eta^{B|A=2}\tau^{B=1}\ \tau^{B=1\ A=2} = \omega^{B|A=2}\ \vartheta^{B=1}\vartheta^{B=1\ A=2}$ (3.7)

ricordando che le costanti di normalizzazione sono:

$$\eta^{B|A=2} = \frac{1}{1+\tau^{B=2}\tau^{B=2\ A=2}}$$

$$\omega^{B|A} = \frac{1}{\vartheta^{B=1}\vartheta^{B=1\ A=2} + \vartheta^{B=2}\vartheta^{B=2\ A=2}} = \frac{\vartheta^{B=1}\vartheta^{B=1\ A=2}}{1+(\vartheta^{B=1}\vartheta^{B=1\ A=2})^2}$$

la (3.7) diviene

$$\frac{1}{1+\tau^{B=2}\tau^{B=2\ A=2}} = \frac{\vartheta^{B=1}\vartheta^{B=1\ A=2}}{1+(\vartheta^{B=1}\vartheta^{B=1\ A=2})^2}\ \vartheta^{B=1}\vartheta^{B=1\ A=2}$$

e sviluppando risulta

$$1+\tau^{B=2}\tau^{B=2\ A=2} = \frac{1+(\vartheta^{B=1}\vartheta^{B=1\ A=2})^2}{(\vartheta^{B=1}\vartheta^{B=1\ A=2})^2} = 1 + \frac{1}{(\vartheta^{B=1}\vartheta^{B=1\ A=2})^2}$$

Il parametro $\tau^{B=2\ A=2}$ è

$$\tau^{B=2\;A=2} = \frac{1}{\tau^{B=2}(\vartheta^{B=1}\vartheta^{B=1\;A=2})^2}$$

- ***parametro*** $\boldsymbol{\tau^{C=2}}$

Per calcolare il parametro $\tau^{C=2}$ si parte dalla probabilità P(C=1|A=1 B=1)

P(C=1|A=1 B=1)= $\eta^{C|A=1\;B=1}\tau^{C=1}\,\tau^{C=1\;A=1}\tau^{C=1\;B=1} = \omega^{C|A=1\;B=1}\vartheta^{C=1}\vartheta^{C=1\;A=1}\vartheta^{C=1\;B=1}$ (3.8)

ricordando che le costanti di normalizzazione sono:

$$\eta^{C|A=1\;B=1} = \frac{1}{1+\tau^{C=2}}$$

$$\omega^{C|A=1\;B=1} = \frac{1}{\vartheta^{C=1}\vartheta^{C=1\;A=1}\vartheta^{C=1\;B=1} + \vartheta^{C=2}\vartheta^{C=2\;A=1}\vartheta^{C=2\;B=1}} = \frac{\vartheta^{C=1}\vartheta^{C=1\;A=1}\vartheta^{C=1\;B=1}}{1+(\vartheta^{C=1}\vartheta^{C=1\;A=1}\vartheta^{C=1\;B=1})^2}$$

la (3.8) diviene

$$\frac{1}{1+\tau^{C=2}} = \frac{\vartheta^{C=1}\vartheta^{C=1\;A=1}\vartheta^{C=1\;B=1}}{1+(\vartheta^{C=1}\vartheta^{C=1\;A=1}\vartheta^{C=1\;B=1})^2}\;\vartheta^{C=1}\vartheta^{C=1\;A=1}\vartheta^{C=1\;B=1}$$

e sviluppando risulta

$$1+\tau^{C=2} = \frac{1+(\vartheta^{C=1}\vartheta^{C=1\;A=1}\vartheta^{C=1\;B=1})^2}{(\vartheta^{C=1}\vartheta^{C=1\;A=1}\vartheta^{C=1\;B=1})^2} = 1+\frac{1}{(\vartheta^{C=1}\vartheta^{C=1\;A=1}\vartheta^{C=1\;B=1})^2}$$

Il parametro $\tau^{C=2}$ è

$$\tau^{C=2} = \frac{1}{(\vartheta^{C=1}\vartheta^{C=1\;A=1}\vartheta^{C=1\;B=1})^2}$$

- ***parametro*** $\boldsymbol{\tau^{C=2\;A=2}}$

Per calcolare il parametro $\tau^{C=2\;A=2}$ si parte dalla probabilità P(C=1|A=2 B=1)

P(C=1|A=2 B=1)= $\eta^{C|A=2\;B=1}\tau^{C=1}\,\tau^{C=1\;A=2}\tau^{C=1\;B=1} = \omega^{C|A=2\;B=1}\vartheta^{C=1}\vartheta^{C=1\;A=2}\vartheta^{C=1\;B=1}$ (3.9)

ricordando che le costanti di normalizzazione sono:

$$\eta^{C|A=2\;B=1} = \frac{1}{1+\tau^{C=2}\tau^{C=2\;A=2}}$$

$$\omega^{C|A=2\;B=1} = \frac{1}{\vartheta^{C=1}\vartheta^{C=1\;A=2}\vartheta^{C=1\;B=1} + \vartheta^{C=2}\vartheta^{C=2\;A=2}\vartheta^{C=2\;B=1}} = \frac{\vartheta^{C=1}\vartheta^{C=1\;A=2}\vartheta^{C=1\;B=1}}{1+(\vartheta^{C=1}\vartheta^{C=1\;A=2}\vartheta^{C=1\;B=1})^2}$$

la (3.9) diviene

$$\frac{1}{1+\tau^{C=2}\tau^{C=2\ A=2}} = \frac{\vartheta^{C=1}\vartheta^{C=1\ A=2}\vartheta^{C=1\ B=1}}{1+(\vartheta^{C=1}\vartheta^{C=1\ A=2}\vartheta^{C=1\ B=1})^2}\ \vartheta^{C=1}\vartheta^{C=1\ A=2}\vartheta^{C=1\ B=1}$$

e sviluppando risulta

$$1+\tau^{C=2}\tau^{C=2\ A=2} = \frac{1+(\vartheta^{C=1}\vartheta^{C=1\ A=2}\vartheta^{C=1\ B=1})^2}{(\vartheta^{C=1}\vartheta^{C=1\ A=2}\vartheta^{C=1\ B=1})^2} = 1+\frac{1}{(\vartheta^{C=1}\vartheta^{C=1\ A=2}\vartheta^{C=1\ B=1})^2}$$

Il parametro $\tau^{C=2\ A=2}$ è

$$\tau^{C=2\ A=2} = \frac{1}{\tau^{C=2}(\vartheta^{C=1}\vartheta^{C=1\ A=2}\vartheta^{C=1\ B=1})^2}$$

- $\boldsymbol{parametro\ \tau^{C=2\ B=2}}$

Per calcolare il parametro $\tau^{C=2\ B=2}$ si parte dalla probabilità P(C=1|A=1 B=2)

P(C=1|A=1 B=2)=$\eta^{C|A=1\ B=2}\tau^{C=1}\ \tau^{C=1\ A=1}\tau^{C=1\ B=2} = \omega^{C|A=1\ B=2}\vartheta^{C=1}\vartheta^{C=1\ A=1}\vartheta^{C=1\ B=2}$ (3.10)

ricordando che le costanti di normalizzazione sono:

$$\eta^{C|A=1\ B=2} = \frac{1}{1+\tau^{C=2}\tau^{C=2\ B=2}}$$

$$\omega^{C|A=1\ B=2} = \frac{1}{\vartheta^{C=1}\vartheta^{C=1\ A=1}\vartheta^{C=1\ B=2}+\vartheta^{C=2}\vartheta^{C=2\ A=1}\vartheta^{C=2\ B=2}} = \frac{\vartheta^{C=1}\vartheta^{C=1\ A=1}\vartheta^{C=1\ B=2}}{1+(\vartheta^{C=1}\vartheta^{C=1\ A=1}\vartheta^{C=1\ B=2})^2}$$

la (3.10) diviene

$$\frac{1}{1+\tau^{C=2}\tau^{C=2\ B=2}} = \frac{\vartheta^{C=1}\vartheta^{C=1\ A=1}\vartheta^{C=1\ B=2}}{1+(\vartheta^{C=1}\vartheta^{C=1\ A=1}\vartheta^{C=1\ B=2})^2}\ \vartheta^{C=1}\vartheta^{C=1\ A=1}\vartheta^{C=1\ B=2}$$

e sviluppando risulta

$$1+\tau^{C=2}\tau^{C=2\ B=2} = \frac{1+(\vartheta^{C=1}\vartheta^{C=1\ A=1}\vartheta^{C=1\ B=2})^2}{(\vartheta^{C=1}\vartheta^{C=1\ A=1}\vartheta^{C=1\ B=2})^2} = 1+\frac{1}{(\vartheta^{C=1}\vartheta^{C=1\ A=1}\vartheta^{C=1\ B=2})^2}$$

Il parametro $\tau^{C=2\ B=2}$ è

$$\tau^{C=2\ B=2} = \frac{1}{\tau^{C=2}(\vartheta^{C=1}\vartheta^{C=1\ A=1}\vartheta^{C=1\ B=2})^2}$$

A.6.1 Esempio con dati del capitolo 1

Si considerano i dati della tabella 14 del capitolo 1. Essi sono costruiti con i seguenti parametri "effect code" :

$$\vartheta^{A} = 0.25$$

$$\vartheta^{B} = 0.8$$

$$\vartheta^{C} = 0.2$$

$$\vartheta^{AB} = 0.5$$

$$\vartheta^{AC} = 1.6$$

$$\vartheta^{CB} = 3.2$$

che possono essere trasformati in parametri dummy code:

$$\tau^{A=2} = \frac{1}{(\vartheta^{A=1})^2} = \frac{1}{(0.25)^2} = 16$$

$$\tau^{B=2} = \frac{1}{(\vartheta^{B=1}\vartheta^{B=1\ A=1})^2} = \frac{1}{(0.8\ 0.5)^2} = 6.25$$

$$\tau^{C=2} = \frac{1}{(\vartheta^{C=1}\vartheta^{C=1\ A=1}\vartheta^{C=1\ B=1})^2} = \frac{1}{(0.2\ 1.6\ 3.2)^2} = 0.9537$$

$$\tau^{B=2\ A=2} = \frac{1}{\tau^{B=2}(\vartheta^{B=1}\vartheta^{B=1\ A=2})^2} = \frac{1}{6.25\left(0.8\ \frac{1}{0.5}\right)^2} = 0.0625$$

$$\tau^{C=2\ A=2} = \frac{1}{\tau^{C=2}(\vartheta^{C=1}\vartheta^{C=1\ A=2}\vartheta^{C=1\ B=1})^2} = \frac{(0.2\ 1.6\ 3.2)^2}{\left(0.2\ \frac{1}{1.6}\ 3.2\right)^2} = 6.5536$$

$$\tau^{C=2\ B=2} = \frac{1}{\tau^{C=2}(\vartheta^{C=1}\vartheta^{C=1\ A=1}\vartheta^{C=1\ B=2})^2} = \frac{(0.2\ 1.6\ 3.2)^2}{\left(0.2\ \frac{1}{3.2}\ 1.6\right)^2} = 104.8576$$

La stima dei dati della tabella 14 (capitolo 1) con l'effect code è riportata nella tabella 15 (capitolo 1), mentre quella con la parametrizzazione dummy code è riportata nella tabella 1 seguente

Tabella 1:stima con dummy code dati tabella 14 cap 1

```
* TABLE A [or P(A)] *

  effect          lambda  exp(lambda)
  A
   1              0.0000   1.0000
   2              2.7726  16.0000

* TABLE AB [or P(B|A)] *

  effect          lambda  exp(lambda)
  B
   1              0.0000   1.0000
   2              1.8326   6.2500
  AB
   1 1            0.0000   1.0000
   1 2            0.0000   1.0000
   2 1            0.0000   1.0000
   2 2           -2.7726   0.0625

* TABLE CAB [or P(C|AB)] *

  effect          lambda  exp(lambda)
  C
   1              0.0000  1.0000
   2             -0.0474  0.9537
  CA
   1 1            0.0000  1.0000
   1 2            0.0000  1.0000
   2 1            0.0000  1.0000
   2 2            1.8800  6.5536
  CB
   1 1            0.0000  1.0000
   1 2            0.0000  1.0000
   2 1            0.0000  1.0000
   2 2            4.6526  104.8576
  CAB
   1 1 1          0.0000  1.0000
   1 1 2          0.0000  1.0000
   1 2 1          0.0000  1.0000
   1 2 2          0.0000  1.0000
   2 1 1          0.0000  1.0000
   2 1 2          0.0000  1.0000
   2 2 1          0.0000  1.0000
   2 2 2          0.0000  1.0000
```

A.7 Effetti diretti- indiretti e totali

Come spiegato nel capitolo 1, la parametrizzazione dummy code porta a delle semplificazioni nella scomposizione dell'effetto totale in diretto, indiretto totale e quest'ultimo in indiretto e cella. Nei capitoli 3 e 4 si è scelto di stimare rispettivamente i dati creati per il confronto tra il modello log-lineare causale e il SEM e quelli del questionario sul prosciutto di Sauris con la parametrizzazione effect code perché, in generale, è la più usata. In questo paragrafo si studiano gli effetti diretti, indiretti e totali con una parametrizzazione effect code. Nei paragrafi precedenti di questa Appendice si evidenzia che mentre nel dummy code l'effetto totale è uguale a $\mu_{dummy\ code}^{C=2\ A=2}$ e l'effetto diretto è uguale a $\tau^{C=2\,A=2}$, nell'effect code l'effetto totale è $\left(\mu_{effect\,code}^{C=2\,A=2}\right)^2$ e l'effetto diretto è $(\vartheta^{AC})^2$. Per semplicità, come spiegato nell'appendice del capitolo 1 e come è stato fatto da Croon et al. (2009), l'effetto diretto si definisce ϑ^{AC} anche nel caso effect code. Di conseguenza si pone l'effetto totale uguale a $\mu_{effect\,code}^{C=2\,A=2}$.

- **Effetto totale**

$$\mu_{dummy\,code}^{C=2\,A=2} = \left(\mu_{effect\,code}^{C=2\,A=2}\right)^4$$

- **Effetto totale scomposto**

Ricordando la formula dell'effetto totale con parametrizzazione dummy code:

$$\mu_{dummy\,code}^{C=2\,A=2} = \tau^{A=2C=2}\left\{\frac{\eta^{C|A=1B=1} + \eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=2B=1} + \eta^{C|A=2B=2}\tau^{B=2}\tau^{B=2A=2}}\left[\frac{\eta^{C|A=1B=1} + \eta^{C|A=1B=2}\tau^{B=2}\tau^{B=2C=2}}{\eta^{C|A=2B=1} + \eta^{C|A=2B=2}\tau^{C=2\,B=2}\tau^{B=2}\tau^{B=2\,A=2}}\right]^{-1}\right\}$$

e sostituendo in essa i parametri effect code

$$\tau^{A=2C=2} = (\vartheta^{A=2C=2})^4$$

$$\tau^{B=2} = \frac{1}{(\vartheta^{B=1}\vartheta^{B=1\,A=1})^2}$$

$$\tau^{B=2C=2} = (\vartheta^{B=2C=2})^4$$

$$\tau^{B=2\,A=2} = (\vartheta^{B=1A=1})^4$$

$$\eta^{C|A=1\,B=1} = \omega^{C|A=1\,B=1}\vartheta^{C=1}\vartheta^{C=1\,A=1}\vartheta^{C=1\,B=1}$$

$$\eta^{C|A=2\,B=1} = \omega^{C|A=2\,B=1}\vartheta^{C=1}\vartheta^{C=1\,A=2}\vartheta^{C=1\,B=1}$$

$$\eta^{C|A=1\,B=2} = \omega^{C|A=1\,B=2}\vartheta^{C=1}\vartheta^{C=1\,A=1}\vartheta^{C=1\,B=2}$$

$$\eta^{C|A=2\,B=2} = \omega^{C|A=2\,B=2}\vartheta^{C=1}\vartheta^{C=1\,A=2}\vartheta^{C=1\,B=2}$$

risulta:

$$\left(\mu_{effect\,code}^{C=2\,A=2}\right)^4 = (\vartheta^{B=2A=2})^4 \left\{\frac{\omega^{C|A=1B=1}\vartheta^{C=1\,B=1} + \frac{\omega^{C|A=1B=2}\vartheta^{C=1\,B=2}}{(\vartheta^{B=1}\vartheta^{B=1\,A=1})^2}}{\omega^{C|A=2B=1}\vartheta^{C=1\,B=1} + \omega^{C|A=2B=2}\frac{(\vartheta^{B=1A=1})^4\vartheta^{C=1\,B=2}}{(\vartheta^{B=1}\vartheta^{B=1\,A=1})^2}} \left[\frac{\omega^{C|A=1B=1}\vartheta^{C=1\,B=1} + \omega^{C|A=1B=2}\frac{(\vartheta^{B=2C=2})^4\vartheta^{C=1\,B=2}}{(\vartheta^{B=1}\vartheta^{B=1\,A=1})^2}}{\omega^{C|A=2B=1}\vartheta^{C=1\,B=1} + \omega^{C|A=2B=2}\frac{\vartheta^{C=1\,B=2}(\vartheta^{B=1C=1})^4(\vartheta^{B=1A=1})^4}{(\vartheta^{B=1}\vartheta^{B=1\,A=1})^2}}\right]^{-1}\right\}$$

semplificando l’equazione si ottiene la scomposizione dell’effetto totale in effetto diretto e indiretto totale con la parametrizzazione effect code:

$$\left(\mu_{effect\ code}^{C=2\ A=2}\right) = \left(\vartheta^{B=2A=2}\right)$$
$$\left\{\frac{\omega^{C|A=1B=1}\vartheta^{C=1\ B=1} + \frac{\omega^{C|A=1B=2}\vartheta^{C=1\ B=2}}{\left(\vartheta^{B=1}\vartheta^{B=1\ A=1}\right)^2}}{\omega^{C|A=2B=1}\vartheta^{C=1\ B=1} + \omega^{C|A=2B=2}\frac{\left(\vartheta^{B=1A=1}\right)^2\vartheta^{C=1\ B=2}}{\left(\vartheta^{B=1}\right)^2}}\right.$$
$$\left.\left[\frac{\omega^{C|A=1B=1}\vartheta^{C=1\ B=1} + \omega^{C|A=1B=2}\frac{\left(\vartheta^{B=2C=2}\right)^4\vartheta^{C=1\ B=2}}{\left(\vartheta^{B=1}\vartheta^{B=1\ A=1}\right)^2}}{\omega^{C|A=2B=1}\vartheta^{C=1\ B=1} + \omega^{C|A=2B=2}\frac{\left(\vartheta^{B=1C=1}\right)^4\left(\vartheta^{B=1A=1}\right)^2\vartheta^{C=1\ B=2}}{\left(\vartheta^{B=1}\right)^2}}\right]^{-1}\right\}^{\frac{1}{4}} \quad (3.11)$$

dove il primo moltiplicatore è l’effetto diretto e il secondo moltiplicatore è l’effetto indiretto totale

- **Scomposizione effetto indiretto totale**

Ricordando come si è ottenuta la formula dell’effetto totale con parametrizzazione dummy code:

$$\text{effetto\_cella}_{dummy\ code}=\frac{\eta^{C|A=1B=1}+\eta^{C|A=1B=2}\tau^{B=2}}{\eta^{C|A=1B=1}+\eta^{C|A=1B=2}\tau^{B=2}\tau^{B=2C=2}}\ \frac{\eta^{C|A=2B=1}+\eta^{C|A=2B=2}\tau^{B=2}\tau^{B=2C=2}}{\eta^{C|A=2B=1}+\eta^{C|A=2B=2}\tau^{B=2}}$$

si ottiene la medesima formula per l’effect code ponendo $\vartheta^{B=1A=1}=1$ nel secondo addendo (effetto indiretto totale) della (3.11)

$$\text{effetto\_cella}_{effect\ code} = \sqrt[4]{\frac{\omega^{C|A=1\ B=1}\vartheta^{C=1\ B=1} + \frac{\omega^{C|A=1\ B=2}\vartheta^{C=1\ B=2}}{\left(\vartheta^{B=1}\right)^2}}{\omega^{C|A=1\ B=1}\vartheta^{C=1\ B=1} + \frac{\omega^{C|A=1\ B=2}\vartheta^{C=1\ B=2}\left(\vartheta^{B=2C=2}\right)^4}{\left(\vartheta^{B=1}\right)^2}}\frac{\omega^{C|A=2\ B=1}\vartheta^{C=1\ B=1} + \frac{\omega^{C|A=2\ B=2}\vartheta^{C=1\ B=2}\left(\vartheta^{B=2C=2}\right)^4}{\left(\vartheta^{B=1}\right)^2}}{\omega^{C|A=2\ B=1}\vartheta^{C=1\ B=1} + \frac{\omega^{C|A=2\ B=2}\left(\vartheta^{C=1\ B=1}\right)}{\left(\vartheta^{B=1}\right)^2}}}$$

e per determinare l’effetto indiretto si usa la formulazione

effetto_indiretto=effetto_ind-tot/effetto_cella

Appendice B

La stima dei modelli fatta con il programma Lem (Vermunt 1997) è molto versatile; comunque esistono altri programmi, come il Latent Gold (Vermunt 2000), che analizzano altrettanto bene i modelli log-lineari. Si può affermare, in generale, che tutto ciò che in letteratura è stato detto sui modelli log-lineari, può essere implementato nel programma Lem (Vermunt 1997 e successivi). Per semplicità si ricordano due esempi: le procedure di stima e le restrizioni per l'identificazione. Le procedure di stima IPF e Newton-Raphson, citate nella teoria, sono entrambe utilizzabili nel programma Lem. I due metodi, effect code e dummy code, usati nella teoria per identificare i modelli, sono possibili scelte nel programma Lem. L'effect code è dato per default, il dummy code deve essere selezionato attraverso il comando dum.

Bibliografia

Baker D. A, Crompton J. L: “ Quality, satisfaction and behavioral intentions” pag 785-803, Annals of Tourism research (2000)

Bartholomew D. J. e Knott M. : “Latent variable models and factor analysis” London: Arnold (1998)

Bartholomew D. J., Steele F., Moustaki I e Galbraith J. I. : “Analysis of multivariate social science data”, CRC Press, second edition (2008)

Bergsma W., Croon M. e Hagenaars J. A. : “ Marginal models for dependent, clustered and longitudinal categorical data “, New York: Springer (2009)

Bollen K. A. : “Structural equations with latent variables”, New York: Wiley (1989)

Chen C., Tsai D.: “ How destination image and evaluative factors affect behavioral intentions?” pag 1115-1122, Tourism management (2006)

Cronin J. J Jr, Brady M. K. e Hult G. T. M. :”Assessing the effect of quality, value, and customer satisfaction on consumer behavioral intentions in service environments” pag 193-218, Journal of Retailing, 76, (2000)

Dunteman G. H., Kim J. O., Long S. J e Mueller C. . “Factor analysis and related techniques” , London: Sage (1994)

Goodman L.A. :”The analysis of multidimensional contingency tables when some variables are posterior to other: a modified path analysis approach” pag 179-192, Biometrika (1973)

Goodman L. A. : “Exploratory latent structure analysis using both identifiable and unidentifiable models” pag 215-231, Biometrika (1974a)

Goodmam L. A: “The analysis of systems of qualitative variables when some variables are unobservable. Part I- A modified latent structure approach” pag 1179-1259, American journal of Sociology (1974b)

Haberman S J:”Analysis of qualitative data”, New York: New Developments, Accademic Press (1979)

Hagenaars J. A. : “Categorical longitudinal data: log-linear analysis of panel, trend and cohort data” Newbury Park, CA: Sage (1990)

Hagenaars J. A. :”Latent structure models with direct effects between indicators:Local dependence models”, pag 379-405, Sociological Methods and Research (1988)

Hagenaars J. A. :”Loglinear models with latent variables”, Newbury Park:Sage (1993)

Heinen T.: “Latent class and discrete latent trait models: similarities and differences”, Thousand Oakes, CA: Sage publications (1996)

Jöreskog, K. G.:”A general method for estimating a linear structural equation system” pag 85-112, in “Structural Equation Models in the Social Sciences”, New York: A. S. Goldberger e O. D. Ducan eds, Accademic press (1973)

Jöreskog, K. G.:”Structural equation models in the social sciences: specification, estimation and testing” pag 265-287, Application of Statistics, Amsterdam: P.R. Krishnaiah ed. (1977)

Keesling, J. W:”MaximumLikelihood Approaches to Casual Analysis”, Ph.D. dissertation, Department of Education: University of Chicago (1972)

Lawley, D. N. e Maxwell A. E. :”Factor Analysis as a Structural Method”, London: Butterworth (1971)

Magidson J e Vermunt J. K. : “Latent GOLD User's Manual”, Boston: Statistical Innovations Inc. (2000)

Magidson J e Vermunt J. K. :”Latent class factor and cluster models, bi-plots, and related grapfhical displays” pag 223-264, Sociological methodology (2001)

Magidson J e Vermunt J. K: “ Factor analysis with categorical indicators: a comparison between traditional and latent class approaches” technical support, Statistical Innovations (2004)

Oliver, R. L.:”Satisfation: A Behavioral Perspectiveon on the Consumer”, New York: MacGraw-Hill (1997)

Pearl J. : “Direct and indirect effects” pag 411-420 in Proceedings of the seventeenth conference on uncertainty in artificial intelligence (2001)

Pearl J.:”The causal mediation formula-A guide to the assessment of pathways and mechanisms” pag 1-17 technical report (2011)

Pistore, C.:”Promuovere i prodotti locali: un modello strutturale per interpretare il behavioral intention”, Tesi di laurea, Padova: Università degli studi (anno accademico 2009/2010)

Rao P. V., Li H. e Roth J.: ”Recursive path models when both predictor and response variables are categorical” pag 664-675, Journal of statistical theory and practice (2008)

Vermunt J. K. : “ Lem: a general program for the analysis of categorical data: users manual”, Tilburg University (1997)

Wiley, D. E.:”The identification problem for structural equation models with unmeasured variables” pag 69-83, in “Structural Equation Models in the Social Sciences”, New York: A. S. Goldberger e O. D. Ducan eds, Accademic press (1973)

Wright, S: “Correlation and Causation”, pag 557-585, Journal of Agricultural Research (1921)