

Output di Lisrel

I t-value tra -2 e 2 sono candidati per essere eliminati (terza riga dei risultati di LX, LY, BE, GA)

Aggiustamento del modello per far sì che si adatti bene ai dati. Partiamo dal χ^2_{n-1}

$$\chi^2 = (N - 1) * C$$

dove C =criterio

$$C = \ln |\Sigma| - \ln |S| + tr(S\Sigma^{-1}) - J$$

dove \ln è il logaritmo naturale, $|\Sigma|$ è il determinante della matrice Σ , tr è la traccia (somma degli elementi diagonali) e J è il numero totale di variabili.

C è un secondo χ^2 che è un piccolo aggiustamento del primo.

Si confronta il χ^2 calcolato con quello in base alla distribuzione teorica: se χ^2 , rispetto ai gl, è alto c'è qualcosa che manca nel modello (guardare gli indici di modifica che ci dicono di quanto si abbassa il χ^2 se inserissimo quel parametro).

Il valore previsto di un χ^2 con K gl è K .

La differenza $(\chi^2 - K)$ è ciò che c'è in più.

Qual è la soglia massima?

La formula di C dipende dalle differenze tra Σ e S .

Se Σ e S coincidono ($\Sigma = S$) la traccia ($tr(S\Sigma^{-1})$) è uguale a J .

Il χ^2 varia in modo lineare con la grandezza del campione; se il campione cresce la stima dovrebbero essere più precisa, quindi ciò che prima era trascurabile ora diventa significativo.

Un χ^2 è alto anche con residui piccoli se il campione è molto grande, perché ci costringe a tener conto di aspetti che abbiamo deciso a priori di ignorare. Es: con dati ordinali

Il χ^2 è quindi potenzialmente troppo sensibile con N grandi.

Quando N è definito troppo grande?

Non ci sono regole fisse:

- fino a 1000 casi in genere le stime di massima verosimiglianza funzionano

- 2000, N è troppo grande
- tra 1000 e 2000 le stime sono affidabili solo tra 25% e 75% con una precisione tra 1% e 3%

Il problema è che nel mondo il modello lineare tiene solo fino a un certo punto!

La maggior precisione della stima dovuta all'aumento di N non garantisce la bontà del χ^2 : lo sfasamento del criterio dovuto al rumore di fondo viene ampliato in campioni grandi.

Con 1000 casi una discrepanza di 0,4 è significativa.

Aggiungendo parametri addizionali da stimare, il χ^2 si abbassa e diventa accettabile, ma si appesantisce (aumentano le frecce, cioè le spiegazioni) e si viola il *criterio di parsimonia*.

Cioè spiegare la realtà con il minor numero di leggi. Ovvero rendere con poche relazioni una situazione complessa.

Solo un modello semplice è davvero informativo. Un modello "a piatto di spaghetti" è da cestinare!

La parsimonia permette anche di tenere le cose sotto controllo: è più facile eliminare le ipotesi sbagliate alla massima velocità possibile.

Un modello può essere al massimo respinto, mai accettato.

Se un χ^2 è basso, ma il modello è molto complesso, potrebbe non andar bene.

La parsimonia contrasta con la procedura di diminuire il χ^2 .

Più legami ipotizziamo, più migliora il χ^2 , meno il modello è parsimonioso.

Come proteggere la parsimonia del modello?

1. criterio AIC (Akaike Information Criterion):

$\chi^2 + 2t$, con t = numero di parametri da stimare.

Ogni volta che si aggiunge un parametro si alza del doppio ($2t$) ma intanto il χ^2 scende. Tra i diversi modelli si sceglie quello con AIC più basso. Anche con un modello saturo l'AIC ha un suo valore.

2. criterio CAIC (AIC corretto):

$$\chi^2 + [1 + \ln(N)]t$$

E' un criterio più forte del precedente.

Con 400 casi, il χ^2 scendere con un valore di 5 per ogni parametro. Con N molto grandi (8000), il χ^2 scende di 7.

Punisce severamente la mancanza di parsimonia e continua a scendere finché si arriva al modello più scarno possibile.

Quando N è grande, in alternativa al χ^2 si usano insieme CAIC e AIC.

RMR=(Root-Mean-square Residual) Residui tra cov-var riprodotte e osservate.

$$\text{RMR} = \sqrt{\frac{2 \sum (s_{ij} - \hat{\sigma}_{ij})^2}{o(o+1)}}, \text{ con } o = \text{osservate}$$

Partendo da var-cov non c'è un limite fisso per RMR. Se si parte da una matrice di correlazione, RMR può variare da -1 a 2.

Il limite da accettare è 0.05. RMR=0.05 significa che la prima cifra delle correlazioni è riprodotta correttamente, mentre non si può dire nulla sulla seconda.

Il limite dell'RMR accettabile dipende anche dallo stato dell'arte.

Nei risultati di Lisrel ci sono anche:

- *AIC saturated model*: diviso 2 da il numero di var/cor ($\text{AIC}/2 = o(o+1)/2$)
- *Model AIC*: si può usare per fare scelte di parsimonia
- *AIC independent model*: indica se ci sono relazioni tra i dati, quando ci sono relazioni molto alte
- *CAIC saturated model*: se *independent* è vicino a *saturated* c'è poca relazione da elaborare. Si vuole in genere che AIC e CAIC del modello siano molto più bassi dei *saturated*

Nell'output di Lisrel si trovano 2 χ^2 :

1. χ^2 del criterio
2. Normal Theory Weighted χ^2 : il criterio sovrastima il χ^2 se i dati sono normali. E' il secondo ad essere usato nell'AIC e CAIC anche se normalmente i due χ^2 sono simili.

Conoscendo il numero di casi, il χ^2 e i gl si può sempre calcolare AIC e CAIC. Il viceversa è molto più complicato.

$$t = \frac{AIC - \chi^2}{2} \quad \chi^2 = AIC - 2t$$
$$t = \frac{CAIC - \chi^2}{1 + \ln(N)} \quad \chi^2 = CAIC - [1 + \ln(N)]t$$

Critical N: dice quanto dovrebbe essere grande N per poter respingere il modello con probabilità 1%.

Lo possiamo usare come indice di bontà.

Dà informazione utile per future ricerche: quanti casi devo avere per poter dimostrare che la ricerca di un collega è sbagliata?

NCP (Estimated non-centrality parameter) Parametro della non centralità e RMSEA

Σ_{pop} e S_{oss} dovrebbero avvicinarsi man mano che N aumenta. Il problema è quanto Σ_{pop} possa essere approssimata a un modello lineare.

$\Sigma_{pop} - \Sigma_{mod}$ è un'approssimazione fissa, anche se non la conosciamo

Σ_{sti} è stimata in base a S_{oss} e dovrebbe avvicinarsi a Σ_{mod} , che è ciò che riusciamo a stimare tramite il modello.

La discrepanza fra Σ_{sti} e S_{oss} viene stimata tramite il χ^2 ed è una discrepanza che non sparirà mai per varie ragioni: non linearità, distribuzione non perfettamente normale...

Come stimare l'errore di approssimazione?

$$\Sigma_{sti} - \Sigma_{oss} \rightarrow \chi^2_{sti}$$

la differenza $\chi^2_{sti} - df$ è ciò che c'è di troppo rispetto all'assunto di causalità lineare: va accreditato alla differenza tra Σ_{pop} e Σ_{mod} .
Es. $70,3 - 40 = 30,2$ dovuto a una non perfetta approssimazione

$$\chi^2_{teorico} = \sum_k z_k^2$$

quando si campiona da normale standardizzata (media 0, dev.st. 1)

E se si campiona da una normale $(\mu, 1)$?

$$\chi^2 = k\mu^2 + \sum_k z_k^2$$

La costante $k\mu^2$ è detta non centralità ed è sempre positiva. Se NC è 0 non dobbiamo preoccuparci perché cade nelle cose che possiamo accettare. Se è maggiore di 0, ci sono problemi.

Il programma stampa anche l'intervallo di confidenza al 90% per questo parametro: deve essere da 0 in poi.

$$RMSEA = \sqrt{\frac{\chi^2 - df}{(N - 1)df}}$$

Dovrebbe stimare $\Sigma_{pop} - \Sigma_{mod}$: differenza tra ciò che è vero e ciò che il modello lineare può fare (ma standardizzata, quindi non importa se si parte da covarianze o da correlazione).

Ha una distribuzione teoricamente valutabile. Si accettano valori 0.05-0.08

Se RMSEA è troppo grande il modello non va bene, se è 0, ci sono legami che possono essere tolti. In un modello saturo si aggiustano errori di approssimazione che non si dovrebbero aggiustare (il modello contiene troppe frecce=robaccia!).

Indici che non si usano: GFI e AGFI

GFI (Goodness-of-fit indices) E' rapporto tra varianza resa dal modello e situazione di indipendenza ma non si conosce la distribuzione teorica di questo indice. In teoria se non è almeno 0.97 o 0.98 il modello non va bene. In realtà il valore è sempre vicinissimo a 1, perciò è difficile valutare la differenza tra modelli:

indica solo che il modello è un minimo ragionevole; in pratica, non si usa.

AGFI=GFI aggiustato in base al numero di parametri che ci sono nel modello. Lo stesso discorso di AGFI: cestinare!

Tutto il resto della sfilata degli indici non serve a molto.

Per riassumere:

Si guarda sempre χ^2 : se va bene (non significativo) finisce lì, altrimenti si guardano gli altri indici [vale se N è molto grande e il χ^2 è significativo].

Indici di modifica

Vengono forniti in automatico da Lisrel, ma solo per le variabili incluse nel modello.

Andrebbero considerati e aggiunti solo gli indici di modifica relativi ai parametri di regressione (LX, LY, GA e BE).

Gli indici di modifica relativi agli errori andrebbero “giustificati”. Sono leciti:

- Con misure ripetute nel tempo. In questo caso la covarianza fra gli errori delle variabili ripetute, può essere interpretato come:
 1. semplice errore (generalmente è isolato fra 2 variabili). Lo si aggiunge
 2. errore sistematico di quella specifica variabile (generalmente è legato alle sole variabili ripetute). Dovrebbe diventare una latente che influenza le misure ripetute di quella variabile
- parole-stimolo. Classico effetto “valanga” di una domanda (in questionario di opinioni o di atteggiamenti) che scatena emozioni o opinioni pregnanti. Non è sempre prevedibile e spesso si scopre a posteriori (particolare momento in cui è stato somministrato il questionario).

Quasi sempre è meglio considerare le covarianze fra errori come “suggerimenti” per l'esistenza di una variabile latente che non è stata presa in considerazione.

Ripensare tutto

Se il modello non funziona, funziona ma è troppo complesso o non è interpretabile in modo logico, conviene buttare tutto e ripartire dalla teoria (analisi della letteratura) per capire cosa si è spagliato nelle ipotesi. Su questa base si cerca di riformulare un modello.

Si può partire dalle sole variabili osservate, e studiare i residui (o gli errori) per identificare le latenti.