PSICOMETRIA ESAME 23 febbraio 2004 ore 12.30-14.30
Testo e correzione

Si ricorda agli studenti di scrivere $\boxed{1}$ **cognome**, $\boxed{2}$ **nome**, $\boxed{3}$ **matricola** e $\boxed{4}$ **corso di laurea** [ST, SC o VO] IN STAMPATELLO su tutti i fogli usati e DI INDICARE CHIARAMENTE A QUALE DOMANDA STANNO RISPONDENDO.

Tutte le risposte sono da riportare sul foglio a protocollo; risposte sul foglio delle domande (che potete tenere e portare a casa) non saranno considerate nella valutazione del compito.

Qualunque forma di copiatura, consultazione, contatto verbale o scritto ed anche il sospetto di queste attività porta all'immediata esclusione dall'esame.

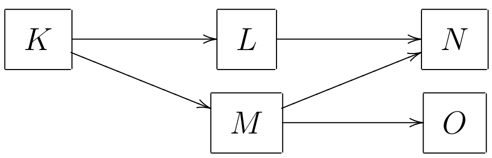
REGRESSIONE

Disponiamo di 5 variabili osservate (K, L, M, N, e O) su un campione di N=200, di cui la tabella che segue mostra le correlazioni. K ha influenza su L e su M; L influenza N; O è influenzata da M; e N è influenzata da L e M.

	K	L	M	N	O
K	1				
L	.35	1			
M	.75	.20	1		
N	.31	.84	.23	1	
O	.17	.28	.68	.37	1

R1. **Disegna il modello causale completo (con tutte le relazioni).**

Ricordiamo che "influenza" equivale ad una freccia orientata. Quindi



volendo, si possono indicare gli errori per L, M, N e O. Ma non per K che è l'unica indipendente. Si possono anche indicare i valori di alcuni parametri (L, M e O) che essendo regressioni semplici (se standardizzati) coincidono con la matrice di correlazione. Non si possono indicare i valori dei parametri di N perché è una regressione multipla e i parametri standardizzati NON coincidono con le correlazioni.

R2. **Scrivi ciascuna delle rette di regressione implicate dal modello (parametri NON standardizzati) usando $a_0 \ldots a_n$ per la prima equazione, $b_0 \cdots b_n$ per la seconda e così via con le altre lettere dell'alfabeto.**

Ogni variabile che riceve una freccia è la dipendente di una equazione di regressione, mentre le variabili che le inviano una freccia sono le indipendenti:

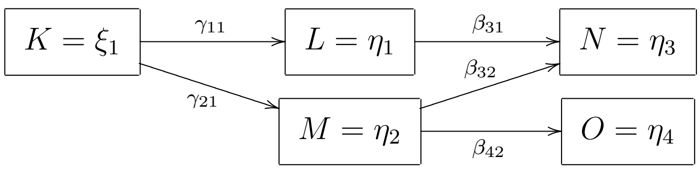
$$\begin{aligned}
L &= a_0 + a_1 K + e \\
M &= b_0 + b_1 K + e \\
N &= c_0 + c_1 L + c_2 M + e \\
O &= d_0 + d_1 M + e
\end{aligned}$$

R3. **Calcola i parametri standardizzati dell'equazione che spiega M e la sua proporzione di varianza spiegata.**

Si fa riferimento all'equazione $M = b_0 + b_1 K$. Si tratta di una regressione semplice e perciò il suo parametro b_1 standardizzato corrisponde alla correlazione fra K e M ($b_1^* = r = .75$ direttamente dalla tabella) e la proporzione di varianza spiegata è il suo quadrato ($r^2 = .75^2 = .56$). $b_0 = 0$ perché la regressione, quando è espressa con dati standardizzati, passa per l'origine.

R4. **Solo VO e SC**: Scrivi il programma Lisrel necessario per testare questo modello (i dati sono nel file "`Test.cor`").

Ricordiamo che, nei modelli con solo variabili osservate, le Y coincidono con le ETA e le X con le KSI. L'unica variabile che non riceve frecce è la K che è quindi una KSI; tutte le altre ricevono e/o mandano frecce, quindi sono delle ETA. I legami fra le ETA sono i BETA e quelli fra ETA e KSI sono i GAMMA. Se completiamo il grafico precedente otteniamo



Vanno poi considerare le varianze delle latenti (PSi e Phi) che però vengono calcolate automaticamente e si potrebbero omettere (le metto tra parentesi quadre). Infine la matrice dati presenta la variabile K come prima variabile in input e quindi dobbiamo usare il comando SE per ordinarle.

```
da ni=5 no=200 ma=km
KM FI="test.cor"
la
K L M N O
SE
2 3 4 5 1 / ! oppure L M N O K /
MO NX=1 NY =4 ga=fu,fi be=fu,fi [ps=di,fr ph=di,fr]
fr ga 1,1 ga 2,1
fr be 3,1 be 3,2 be 4,2
pd
ou
```

R5. **Solo ST: Dai una breve definizione di multicollinearità (max 5 righe)**.

Si parla di multicollinearità quando due o più variabili che vorremmo inserire in una regressione multipla hanno correlazioni molto alte fra di loro.

R6. **Solo ST: Spiega a cosa serve un test di significatività su un parametro di regressione (max 5 righe)**.

Il test di significatività su un parametro di regressione serve per verificare che l'apporto esplicativo di quella variabile sia statisticamente sostanzioso. Ovvero, se quel parametro viene posto a zero (e quindi quella variabile non viene considerata), la retta di regressione restante continua a spiegare la stessa quantità di varianza?

ANALISI FATTORIALE

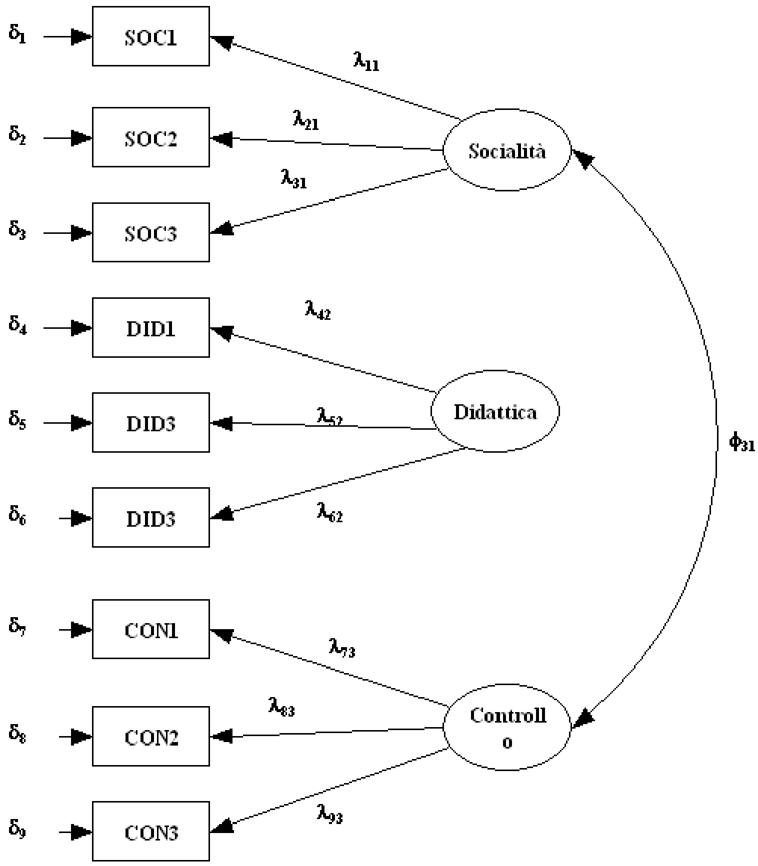
In una ricerca con 78 genitori che hanno avuto da poco il loro primo figlio, è stato usato il questionario SPPQ (Social Project Perspective Questionnaire) di 22 items. In una successiva analisi, i primi 3 item vengono usati per misurare la variabile latente *Socialità*, i successivi 3 item per la variabile latente *Didattica* ed infine altri 3 per la latente *Controllo*.

Solo ST: Socialità e Controllo correlano fra loro, mentre Didattica non correla con nessuno degli altri.

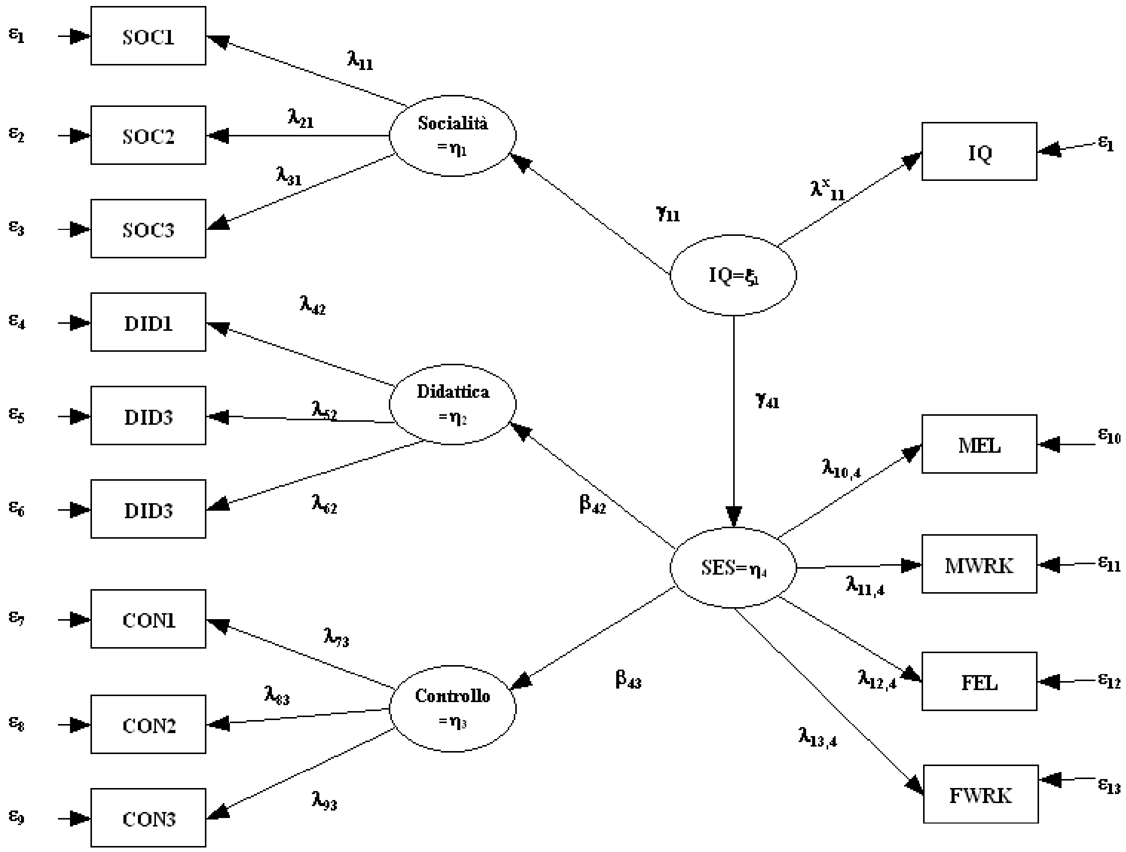
Solo VO e SC: Contemporaneamente è stato misurato l'IQ della madre e il livello educativo e il tipo di lavoro di entrambi i genitori (rispettivamente MEL, MWRK, FEL, FWRK). Livello educativo e tipo di lavoro servono per stimare la variabile latente SES (Socio-economical status). Si ipotizza che l'IQ della madre abbia influenza sulla Socialità e sul SES, mentre il SES ha influenza su Didattica e Controllo.

F1.1. **Disegna il modello causale completo (con tutte le relazioni) indicando eventuali parametri prefissati da te.**

ST: Si tratta di un semplice modello di analisi fattoriale:



VO/SC: Abbiamo un modello di analisi causale con variabili osservate e variabili latenti. Il primo passo è quello di identificare quali osservate servono per misurare quali latenti: dalle istruzioni vediamo che le 3 variabili SOC1-SOC3 misurano la latente Socialità; DID1-DID3 Didattica; CON1-CON3 Controllo; MEL, MWRK, FEL e FWRK misurano la latente SES. L'osservata IQ influenza Socialità e SES, quindi l'osservata IQ deve servire per misurare perfettamente una latente IQ. La latente IQ è l'unica latente che manda frecce ma non ne riceve alcuna, quindi è una KSI. Tutte le altre latenti ricevono almeno una freccia e quindi sono delle ETA. Tutte le osservate (tranne IQ) sono allora delle Y, mentre l'osservata IQ è una X.



F1.2. Scrivi il contenuto di tutte le matrici implicate dal modello, indicando eventuali parametri prefissati da te.

ST: In un modello di analisi fattoriale, le matrici implicate sono solo 3.

$$\Delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \\ \delta_7 \\ \delta_8 \\ \delta_9 \end{bmatrix} \quad \Lambda_x = \begin{bmatrix} \lambda^x_{11} & 0 & 0 \\ \lambda^x_{21} & 0 & 0 \\ \lambda^x_{31} & 0 & 0 \\ 0 & \lambda^x_{42} & 0 \\ 0 & \lambda^x_{52} & 0 \\ 0 & \lambda^x_{62} & 0 \\ 0 & 0 & \lambda^x_{73} \\ 0 & 0 & \lambda^x_{83} \\ 0 & 0 & \lambda^x_{93} \end{bmatrix} \quad \Phi = \begin{bmatrix} 1 & & \\ 0 & 1 & \\ \phi_{31} & 0 & 1 \end{bmatrix}$$

$$\Theta^\delta = \begin{bmatrix} \theta_1^\delta & & & & & & & & \underline{0} \\ & \theta_2^\delta \\ & & \theta_3^\delta \\ & & & \theta_4^\delta \\ & & & & \theta_5^\delta \\ & & & & & \theta_6^\delta \\ & & & & & & \theta_7^\delta \\ & & & & & & & \theta_8^\delta \\ \underline{0} & & & & & & & & \theta_9^\delta \end{bmatrix}$$

VO/SC: In questo modello causale si usano tutte le 8 matrici.

$$\Lambda_y = \begin{bmatrix} \lambda_{11}=1 & 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 & 0 \\ \lambda_{31} & 0 & 0 & 0 \\ 0 & \lambda_{42}=1 & 0 & 0 \\ 0 & \lambda_{52} & 0 & 0 \\ 0 & \lambda_{62} & 0 & 0 \\ 0 & 0 & \lambda_{73}=1 & 0 \\ 0 & 0 & \lambda_{83} & 0 \\ 0 & 0 & \lambda_{93} & 0 \\ 0 & 0 & 0 & \lambda_{10,4}=1 \\ 0 & 0 & 0 & \lambda_{11,4} \\ 0 & 0 & 0 & \lambda_{12,4} \\ 0 & 0 & 0 & \lambda_{13,4} \end{bmatrix} \qquad \Lambda_x = \begin{bmatrix} \lambda_{11}=1 \end{bmatrix} \qquad \Gamma = \begin{bmatrix} \gamma_{11} \\ 0 \\ 0 \\ \gamma_{41} \end{bmatrix}$$

$$B = \begin{bmatrix} 0 \\ 0 & 0 \\ 0 & 0 & 0 \\ 0 & \beta_{42} & \beta_{43} & 0 \end{bmatrix} \qquad \Theta^\epsilon = \begin{bmatrix} \theta_1 & & & & & & & & \underline{0} \\ & \theta_2 \\ & & \theta_3^\epsilon \\ & & & \theta_4^\epsilon \\ & & & & \theta_5^\epsilon \\ & & & & & \theta_6^\epsilon \\ & & & & & & \ddots \\ & & & & & & & \ddots \\ \underline{0} & & & & & & & & \theta_{13}^\epsilon \end{bmatrix}$$

$$\Theta^\delta = \begin{bmatrix} \theta_{11}^\varepsilon \end{bmatrix} \qquad \Phi = \begin{bmatrix} \phi_{11} \end{bmatrix} \qquad \Psi = \begin{bmatrix} \psi_{11} \\ \psi_{21} & \psi_{22} \\ \psi_{31} & \psi_{32} & \psi_{33} \\ \psi_{41} & \psi_{42} & \psi_{43} & \psi_{44} \end{bmatrix}$$

F1.3. **Calcola i gradi di libertà, indicando quali e quanti parametri concorrono a formare la parte *t* della formula.**

ST: Abbiamo 9 λ_x, 9 δ e 1 covarianza (ϕ_{31}), quindi

$$\frac{9(10)}{2} - (9+9+1) = 45 - 19 = 26$$

VO/SC: Abbiamo 9 λ_y, 2 γ, 2 β, 13 θ^ϵ, 1 $\theta\delta$, 1 ϕ e 10 ψ.

$$\frac{14(15)}{2} - (9+2+2+13+1+1+10) = 105 - 38 = 67$$

F1.4. **I dati sono disponibili nel file "`SPPQ.COV`" e le variabili sono presentate nel seguente ordine: SOC1, SOC2, SOC3, DID1, DID2, DID3, CON1, CON2, CON3, IQ, MEL, MWRK, FEL, FWRK. Scrivi il programma Lisrel necessario per stimare questo modello.**

ST: Il file SPPQ.COV contiene 14 osservate e noi usiamo solo le prime 9.

```
da ni=14 no=78 ma=km
la
soc1 soc2 soc3 did1 did2 did3 con1 con2 con3 /
cm fi=SPPQ.COV
se
1 2 3 4 5 6 7 8 9 /
mo nx=9 nk=3 lx=fu,fi ph=sy,fi td=di,fr
lk
social didat contr /
fr lx 1,1 lx 2,1 lx 3,1 lx 4,2 lx 5,2 lx 6,2 lx 7,3 lx 8,3 lx 9,3
fr ph 3,1
va 1 ph 1,1 ph 2,2 ph 3,3 ! default di Lisrel
pd
ou
```

VO e SC: Il file SPPQ.COV contiene 14 osservate, ma la 10ma è una X, le altre delle Y.

```
da ni=14 no=78 ma=km
la
soc1 soc2 soc3 did1 did2 did3 con1 con2 con3 iq mel mwrk fel fwrk
cm fi=SPPQ.COV
se
1 2 3 4 5 6 7 8 9 11 12 13 14 10 /
mo ny=13 nx=1 ne=4 nk=1 ly=fu,fi ga=fu,fi be=fu,fi ps=sy,fr c
  lx=fu,fi ph=sy,fr te=di,fr td=di,fr
le
social didat contr ses /
lk
IQ1
fr ly 2,1 ly 3,1 ly 5,2 ly 6,2 ly 8,3 ly 9,3
fr ly 11,4 ly 12,4 ly 13,4
fr be 2,4 be 3,4
fr ga 1,1 ga 4,1
! le 2 righe seguenti sono di default
```

```
va 1 ly 1,1 ly 4,2 ly 7,3 ly 10,4
va 1 lx 1,1
pd
ou
```

F2.1. **Indica le condizioni generali necessarie per effettuare un'analisi di modelli strutturali in Lisrel (in particolare con l'approccio della massima verosimiglianza).**

Dal momento che si basa tutto su regressioni lineari, le condizioni necessarie sono che le variabili si distribuiscano in modo normale multivariato e che la loro relazione reciproca sia lineare.

F2.2. **Indica almeno due degli assunti addizionali all'interno di un modello di analisi fattoriale confermativa (solo ST) o di un modello causale con variabili latenti (per VO e SC).**

Che gli errori delta non correlino con le ksi; che gli errori non correlano fra loro.

F2.3. **Se Lisrel, ci informa che un determinato modello ha ottenuto i seguenti valori di `Goodness of Fit` [N=2000, χ^2=63.37, df=5, P=.000 RMSEA=.076], quali conclusioni trai sul modello testato? (max 5 righe)**

Il chi quadro è significativo, il che significa che il modello non si adegua a sufficienza ai dati osservati.

F2.4. **Scrivi l'equazione o le equazioni matriciali compatte dei modelli di misura di Lisrel.**

$\mathbf{Y} = \mathbf{\Lambda_y}\eta + \epsilon \qquad \mathbf{X} = \mathbf{\Lambda_x}\xi + \delta$

F2.5. **Se, dopo aver effettuato un'analisi fattoriale confermativa, gli indici di modifica della matrice THETA-DELTA sono maggiori di 4 in tutta l'area inferiore alla diagonale principale, cosa significa?**

Significa che gli errori delta correlano fra loro e questo viola uno degli assunti dell'AF.

F2.6. **Quali possibili azioni intraprendi, dopo F2.5? Elencane almeno una.**

Aggiungere un fattore all'analisi.