ESAME DI PSICOMETRIA - 9 febbraio 2005 - ore 10.30
SOLUZIONI

La barra nera sulla sinistra, indica le mie spiegazioni, per aiutarvi a capire.

Scienze e Tecniche Psicologiche

REGRESSIONE

Usando i dati della tabella seguente (N=10), rispondi alle domande da R1 a R5.

	Y	X_1	X_2	X_3	Medie
Y	9,167				5,50
X_1	8,150	9,833			8,50
X_2	4,750	2,650	20,056		20,50
X_3	-1,550	0,000	-3,950	12,767	5,90

R1 **(2 punti)** Stima tutti i parametri standardizzati dell'eq. $Y = b_0 + b_1X_1 + b_3X_3 + e$

L'equazione è una regressione multipla con due variabili indipendenti fra di loro non correlate (la cov fra X_1 e X_3 è pari a 0), quindi ciascuno dei due parametri b_1 e b_2 può essere calcolato come se fossero parte di una regressione semplice, quindi con la formula della correlazione. Anche b_0 potrebbe essere calcolato perché disponiamo delle medie, ma una volta standardizzato è comunque uguale a 0.

$$b_1^* = \frac{cov(Y, X_1)}{\sqrt{var(Y)var(X_1)}} = \frac{8,150}{\sqrt{9,167 \times 9,833}} = \frac{8,150}{\sqrt{90,139}} = \frac{8,150}{9,494} = 0,858$$
$$b_3^* = \frac{cov(Y, X_3)}{\sqrt{var(Y)var(X_3)}} = \frac{-1,550}{\sqrt{9,167 \times 12,767}} = \frac{-1,550}{\sqrt{117,035}} = \frac{-1,550}{10,818} = -0,143$$
$$b_0^* = 0$$

R2 **(1 punti)** Calcola l'R^2 dell'eq. precedente

Dal momento che, nell'equazione precedente, b^* coincide con r, è sufficiente fare la somma dei quadrati dei $b^* = r$.

$$R^2 = 0,858^2 + (-0,143)^2 = 0,757$$

R3 **(5 punti)** Stima tutti i parametri non standardizzati dell'eq. $Y = c_0 + c_2X_2 + c_3X_3 + e$

Anche in questo caso, si tratta di una una regressione multipla con due variabili indipendenti, ma sono correlate, e quindi ciascuno dei due parametri c_2 e c_3 deve essere calcolato tramite la formula matriciale $\mathbf{C^{-1}c}$. Anche c_0 può essere calcolato perché disponiamo delle medie.

$$\mathbf{C} = \begin{bmatrix} 20,056 & -3,950 \\ -3,950 & 12,767 \end{bmatrix} \qquad \mathbf{c} = \begin{bmatrix} 4,750 \\ -1,550 \end{bmatrix} \qquad \left|\mathbf{C}\right| = 20,056 \times 12,767 - (-3,950)^2$$

$$\frac{1}{240,452} \begin{bmatrix} 12,767 & 3,950 \\ 3,950 & 20,056 \end{bmatrix} \begin{bmatrix} 4,750 \\ -1,550 \end{bmatrix} = \frac{1}{240,452} \begin{bmatrix} 12,767 \times 4,750 + 3,950 \times (-1,550) \\ 3,950 \times 4,750 + 20,056 \times (-1,550) \end{bmatrix} =$$

$$= \begin{bmatrix} 0,227 \\ -0,051 \end{bmatrix} \qquad c_0 = 5,5 - 0,227 \times 20,50 - (-0,051) \times 5,90 = 1,147$$

R4 **(4 punti)** Calcola l'R^2 dell'eq. precedente

> Bisogna prima standardizzare i due parametri c_2 e c_3 e poi calcolare le correlazioni fra X_2, X_3 e Y. Ma si può mettere tutto assieme e semplificare:
>
> $$c_2^* r_{yx_2} = c_2 \frac{\sqrt{var(X_2)}}{\sqrt{var(Y)}} \frac{cov(Y, X_2)}{\sqrt{var(X_2) var(Y)}} = c_2 \frac{cov(Y, X_2)}{var(Y)}$$

$$R^2 = c_2 \frac{cov(Y, X_2)}{var(Y)} + c_3 \frac{cov(Y, X_3)}{var(Y)} = 0,227 \frac{4,75}{9,167} + (-0,051) \frac{-1.55}{9,167} = 0,009$$

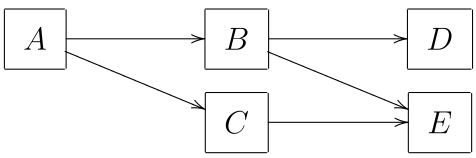
R5 **(3 punti)** Tramite la statistica di F, calcola la significatività del parametro c_2 (senza verificare la significatività)

> Per calcolare F si servono: R^2_f ed R^2_r. Quello full lo abbiamo calcolato al punto precedente, per quello ridotto bisogna formulare l'equazione ridotta, che è $Y = c_0 + 0X_2 + c_3 X_3 + e$ ovvero una regressione semplice. I gradi di libertà saranno N-3 e N-2

$$F = \frac{(R^2_f - R^2_r)/(d_r - d_f)}{(1 - R^2_f)/d_f} = \frac{[0,009 - (-0,143)^2]/(8-7)}{(1 - 0,009)/7} =$$

MODELLO CAUSALE

Considera il seguente modello causale:



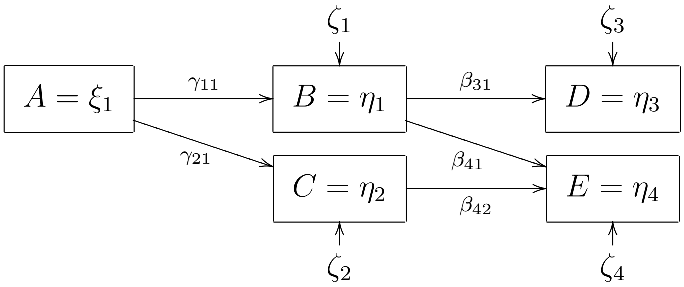
C1 **(1 punti)** Indica quali sono le variabili endogene e quali le esogene

> Le variabili endogene sono quelle interne al modello che vengono spiegate da altre variabili. Quelle esogene sono quelle esterne al modello perché non vengono spiegate.

A è una variabile esogena; B, C, D ed E sono endogene

C2 **(3 punti)** Ri-disegna il modello aggiungendo l'indicazione dei parametri e gli errori (procedi in ordine alfabetico)

Si dovrebbe cominciare identificando le variabili con le ξ o con le η, numerandole, aggiungendo i parametri γ fra KSI ed ETA, i β fra le ETA ed infine gli errori ζ sulle ETA.



C3 **(4 punti)** Scrivi il programma Lisrel (ipotizza N=100, dati nel file XXX.COV)

```
DA NI=5 NO=100 MA=CM
CM FI=XXX.COV
LA; A B C D E
SE; B C D E A
MO NX=1 NY=4 GA=FU,FI, BE=FU,FI PS=DI,FR PH=DI,FR
FR GA 1 1 GA 2 1 BE 3 1 BE 4 1 BE 4 2
PD; OU
```

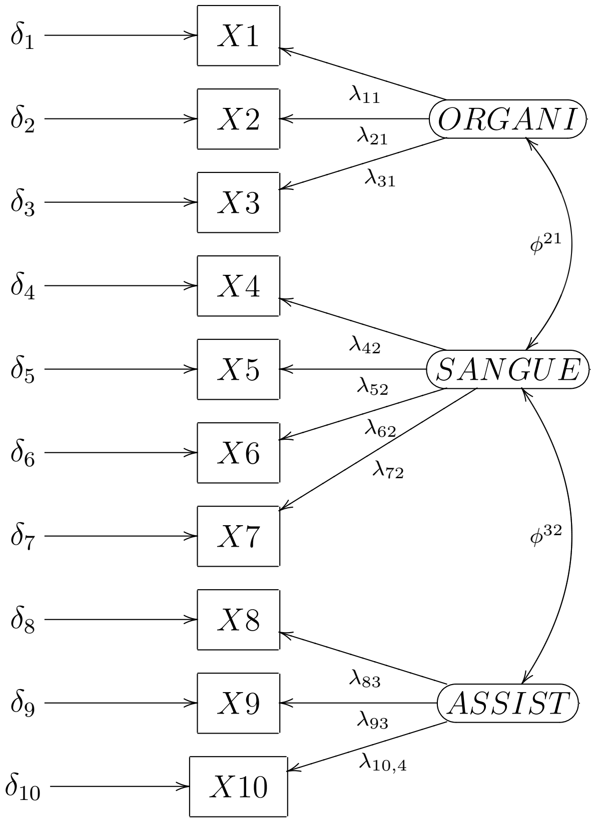
ANALISI FATTORIALE

In uno studio sulla percezione del rischio HIV negli operatori sanitari è stato somministrato ad un gruppo di 98 operatori ospedalieri un questionario per valutare la rischiosità percepita di alcune operazioni di routine (i dati sono raccolti nel file HIV.COR). Dall'analisi fattoriale esplorativa sono emerse tre dimensioni latenti.

La prima dimensione è stata nominata "contatto con materiale e secrezioni organiche" (ORGANI) e raggruppa gli item da 1 a 3. La seconda dimensione è stata nominata "contatto con sangue" (SANGUE) e raggruppa gli item da 4 a 7. La terza dimensione è stata nominata "assistenza al paziente" (ASSIST) e raggruppa gli item da 8 a 10. I risultati hanno chiaramente mostrato che la prima e la seconda dimensione sono fortemente correlate tra loro, mentre la terza dimensione correla in modo debole solo con la seconda.

F1 **(4 punti)** Disegna il grafico causale del modello;

Si tratta di una semplice analisi fattoriale. In aggiunta, sappiamo che il primo fattore correla con il secondo, il secondo con il terzo, ma il terzo non correla con il primo.



F2 **(3 punti)** Scrivi le matrici dei parametri che verranno calcolati;

Essendo un'analisi fattoriale classica e non essendoci motivi particolari per fare altrimenti, parametrizziamo sulle varianze.

$$LX = \Lambda^x = \begin{bmatrix} \lambda_{11} & 0 & 0 \\ \lambda_{21} & 0 & 0 \\ \lambda_{31} & 0 & 0 \\ 0 & \lambda_{42} & 0 \\ 0 & \lambda_{52} & 0 \\ 0 & \lambda_{62} & 0 \\ 0 & \lambda_{72} & 0 \\ 0 & 0 & \lambda_{83} \\ 0 & 0 & \lambda_{93} \\ 0 & 0 & \lambda_{10,3} \end{bmatrix} \quad PH = \Phi = \begin{bmatrix} 1 & & \\ \phi_{21} & 1 & \\ 0 & \phi_{32} & 1 \end{bmatrix} \quad TD = \Theta^\delta = \begin{bmatrix} \theta^\delta_{11} & & \underline{0} \\ & \ddots & \\ \underline{0} & & \theta^\delta_{10,10} \end{bmatrix}$$

F3 **(4 punti)** Scrivi il programma Lisrel.

```
DA NI=10 NO=98 MA=KM
KM=HIV.COR
MO NX=10 NK=3 LX=FU,FI PH=SY,FI TD=DI,DR
FR LX 1,1 LX 2,1 LX 3,1 LX 4,2 LX 5,2 LX 6,2 LX 7,2
FR LX 8,3 LX 9,3 LX 10,3 PH 2,1 PH 3,2
VA 1.0 PH 1,1 PH 2,2 PH 3,3
LK; ORGANI SANGUE ASSIST
PD; OU
```

Considera ora la seguente matrice di saturazioni, ottenuta da un'AFE su un questionario di stili materni.

Matrice fattoriale ruotata Equamax

	Fattore		
	1	2	3
MMA8	0,721	0,366	0,216
MMA10	0,661	-0,020	0,028
MMA9	0,646	0,121	0,258
MMA15	0,530	0,267	0,229
MMA5	0,521	0,382	-0,076
MMA6	0,228	0,078	0,118
MMA7	0,114	0,787	0,039
MMA12	0,164	0,535	0,090
MMA4	0,055	0,507	0,179
MMA2	0,206	0,483	0,159
MMA3	0,308	0,432	0,207
MMA13	0,305	0,346	0,012
MMA1	0,101	0,340	0,168
MMA16	0,181	0,130	0,768
MMA17	0,225	0,223	0,735
MMA14	0,159	0,241	0,564
MMA11	-0,117	-0,099	0,179

F4 **(1 punti)** Calcola l'unicità dell'item MMA5

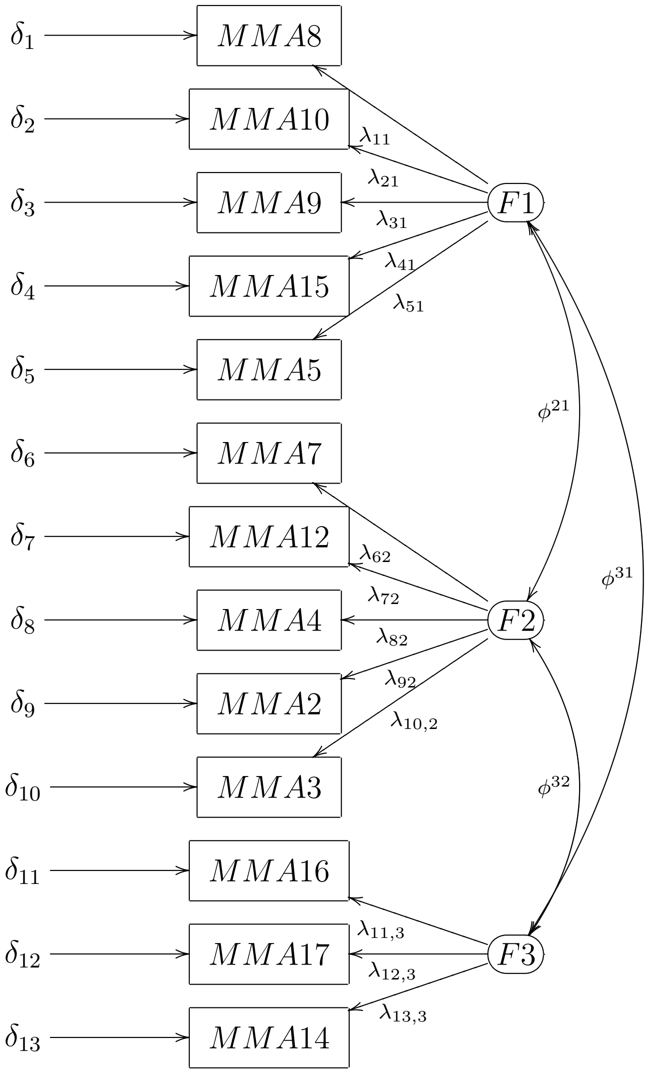
L'unicità è la parte di varianza di un item non spiegata dai fattori, cioè la comunalità. Per poter calcolare la comunalità bisogna usare una soluzione ortogonale. La rotazione Equamax è ortogonale.

$$u^2_{MM5} = 1 - [0,521^2 + 0,382^2 + (-0,076)^2] = 0,577$$

F5 **(7 punti)** Proponi un modello confermativo per questa analisi esplorativa (o tramite un grafico causale o tramite un programma Lisrel).

Da un'analisi fattoriale esplorativa si passa ad una confermativa selezionando gli item più saturi per ciascun fattore (utilizzando livelli superiori a .40 o a .35). In genere si cerca di ottenere un modello congenerico, ovvero un modello in cui ogni item è influenzato da un solo fattore. La matrice presenta già gli item ordinati in modo da facilitare l'interpretazione. Usando il livello di 0.35 avremmo due item bi-fattoriali (MMA8 e MMA5), con il livello .40 invece avremmo un modello congenerico. Gli item 8, 10, 9, 15 e 5 per il fattore 1; 7, 12, 4, 2 e 3 per il fattore 2; 16, 17 e 14 per il fattore 3. 4 item non vengono presi in considerazione.

```
DA NI=17 NO...
...
SE; 8 10 9 15 5 7 12 4 2 3 16 17 14 /
MO NX=13 NK=3 LX=FU,FI PH=ST TD=DI,FR
FR LX 1,1 LX 2,1 LX 3,1 LX 4,1 LX 5,1
FR LX 6,2 LX 7,2 LX 8,2 LX 9,2 LX 10,2
FR LX 11,3 LX 12,3 LX 13,3
PD; OU
```



F6 **(2 punti)** Partendo da un $\chi^2 = 168,31$ (con gl=57 e N=306), calcola e interpreta l'RMSEA del modello

$$RMSEA = \sqrt{\frac{\chi^2 - gl}{N \times gl}} = \sqrt{\frac{168,31 - 57}{306 \times 57}} = 0,08$$

L'RMSEA ottenuto non è buonissimo (dovrebbe essere inferiore a 0,05) ma essendo comunque inferiore a 0,10 è sufficiente per dire che il modello si adegua in modo mediocre ai dati.

F7 **(2 punti)** Se abbiamo $AIC = 236,31$ (usando i dati di F6), qual è il valore di q?

$$AIC = \chi^2 - 2q \quad 236,31 = 168,31 - 2q \quad 236,31 - 168,31 = 2q \quad \frac{68}{2} = q \quad q = 34$$

ESAME DI PSICOMETRIA - 9 febbraio 2005 - ore 10.30
Scienze della Comunicazione - Psicologia (Vecchio Ordinamento)

Le risposte alle domande R1-R5, C1-C3, F1-F3 e F6-F7 sono identiche a quelle precedenti

Ipotizza di avere un secondo campione (file dati HIV2.COR, N=102) proveniente da un'altra azienda ospedaliera e di voler verificare:

F4 **(4 punti)** Che il modello dei lambda precedente si applichi esattamente a questo secondo campione (parametri vincolati). Scrivi i cambiamenti nel programma.

Posso fare questa verifica in due modi: vincolando i parametri LX di questo campione agli stessi valori ottenuti dal modello precedente oppure impostando un modello multisample. Non conoscendo (in questo momento) gli esatti valori di LX, usiamo il secondo metodo.

Alla prima riga del programma aggiungiamo `NG=2`, cioè

```
DA NI=10 NO=98 MA=KM NG=2
```

e dopo il comando `OU`, impostiamo le informazioni per il secondo campione

```
DA NO=102
KM=HIV2.COR
MO LX=IN
PD; OU
```

F5 **(4 punti)** Che nel secondo campione si abbia la stessa struttura fattoriale ma non gli stessi valori. Scrivi i cambiamenti nel programma

Anche in questo caso posso fare questa verifica in due modi: fissando i parametri LX di questo campione nello stesso modo del modello precedente oppure usando un modello multisample.

In questo secondo caso, alla prima riga del programma aggiungiamo `NG=2`, cioè

```
DA NI=10 NO=98 MA=KM NG=2
```

e dopo il comando `OU`, impostiamo le informazioni per il secondo campione

```
DA NO=102
KM=HIV2.COR
MO LX=PS
PD; OU
```