

PSP6075525 - Testing psicologico (matr. dispari)

Caso studio del 08-02-22

Istruzioni iniziali

- Si avvii una nuova sessione di R (o RStudio).
- Si crei un nuovo script di R e lo si salvi come `cognome_nome.R`.
- Si effettui il download del file di dati dell'esame `dati_esame.Rdata` disponibile presso la pagina moodle del corso e lo si carichi nell'ambiente di lavoro di R.
- Si crei un nuovo documento di testo (mediante LibreOffice Writer, Microsoft Word o software analogo) e lo si salvi come `cognome_nome.doc`. Il file dovrà contenere le risposte ai quesiti d'esame accompagnati dai comandi di R, dai risultati ottenuti e dai grafici prodotti. Le risposte dovranno essere inserite in ordine, rispettando il numero del quesito a cui si riferiscono. Alla fine, il file dovrà essere convertito in formato non modificabile (PDF: `cognome_nome.pdf`) ed inviato al docente utilizzando la procedura "Consegna documento" disponibile presso la pagina Moodle del corso. Nel caso di utilizzo di **R-markdown** per la compilazione dinamica di documenti di testo, sarà necessario inviare il file sorgente `.Rmd` unitamente al file PDF generato. Si ricorda di riportare chiaramente Nome, Cognome e Matricola all'interno dei file contenenti le soluzioni finali (`.pdf`, `.R`, `.Rmd`).
- La valutazione della prova sarà effettuata utilizzando primariamente il file `cognome_nome.pdf`: si raccomanda pertanto la chiarezza nella scrittura delle risposte e la correttezza nel riportare i comandi e gli output di R. Il file `cognome_nome.R` dovrà essere allegato al file `cognome_nome.pdf` solo per un controllo aggiuntivo (pertanto non verrà primariamente valutato).

Caso studio

Il caso studio si riferisce alla valutazione del test DASS-21 (*Depression Anxiety Stress Scale*) usato per la valutazione della depressione, ansia e stress. Il test è composto da tre dimensioni (D: depressione; S: stress; A: ansia) quantificate mediante 21 item. I dati si riferiscono ad uno studio¹ che ha coinvolto 738 partecipanti (di cui 374 di genere maschile) di nazionalità britannica. Gli item (frequenza di comportamenti) sono stati rilevati su scale ordinali a 4 livelli (1: “Did not apply”, ..., 4: “Applied most of the time”). Alcuni di questi item sono i seguenti: *How often..Life was meaningless*, *Unable to become enthusiastic*, *Felt scared without reason*, *Difficult to relax*. L'obiettivo dell'analisi è quello di studiare la dimensionalità complessiva del test DASS-21 anche mediante l'ausilio dell'analisi dell'invarianza.

1. Si individuino il numero di unità statistiche e si commenti il tipo di dato a disposizione.

Il numero di unità statistiche è pari a $n = 738$ (maschi: $n = 364$; femmine: $n = 374$). I dati a disposizione consistono nelle risposte su scala ordinale ai $p = 21$ item del test.

2. Si definisca e si adatti ai dati un modello fattoriale confermativo a tre variabili latenti secondo la seguente assegnazione:

D: DASS_Q3 DASS_Q10 DASS_Q13 DASS_Q16 DASS_Q17 DASS_Q21

A: DASS_Q9 DASS_Q15 DASS_Q20

S: DASS_Q6 DASS_Q8 DASS_Q11 DASS_Q12

Il modello è composto da $q = 3$ variabili latenti per $p = 13$ variabili osservate secondo l'equazione lineare

$$\Sigma_{y_{13 \times 13}} = \Lambda_{13 \times 3} \Phi_{3 \times 3} \Lambda_{13 \times 3}^T + \Theta_{\delta_{21 \times 21}}$$

L'adattamento ai dati $\mathbf{S}_{y_{13 \times 13}}$ deve essere fatto mediante stimatori DWLS per dati ordinali. Il modello necessita di 44 parametri da stimare (13 coefficienti fattoriali, 13 varianze d'errore, 2 varianze delle variabili latenti) su un totale di $p(p+1)/2 = 91$ parametri totali. Poiché la stima è effettuata mediante DWLS, vi sono parametri aggiuntivi da stimare (c.d. *thresholds* degli item) e che si riferiscono alle soglie continue associate alle categorie di risposta.

```
#ricodifichiamo ciascuna variabile osservata in termini di 'ordered factor'
for(j in 3:23){
  dass21[,j] = factor(dass21[,j],ordered=TRUE)
}

mD = paste0("D =~ ", paste(paste0("DASS_Q",c(3,10,13,16,17,21)),collapse=" + "))
mA = paste0("A =~ ", paste(paste0("DASS_Q",c(9,15,20)),collapse=" + "))
mS = paste0("S =~ ", paste(paste0("DASS_Q",c(6,8,11,12)),collapse=" + "))

jjd = paste0("DASS_Q",c(3,10,13,16,17,21,9,15,20,6,8,11,12))

model1 = paste(mD,mA,mS,sep = "\n ")
fit1 = lavaan::cfa(model = model1,data = dass21[,jjd],ordered = jjd,estimator="DWLS")
```

3. Si interpreti il risultato del modello adattato al punto precedente anche mediante l'utilizzo di indici di adattamento complessivo (si suggerisce l'utilizzo dei coefficienti standardizzati nell'interpretazione della soluzione fattoriale).

¹ Gomez, R., Stavropoulos, V., & Griffiths, M. D. (2020). Confirmatory factor analysis and exploratory structural equation modelling of the factor structure of the Depression Anxiety and Stress Scales-21. *PloS one*, 15(6), e0233998.

```
print(fit1)
```

lavaan 0.6-7 ended normally after 33 iterations

Estimator	DWLS
Optimization method	NLMINB
Number of free parameters	55
Number of observations	657

Model Test User Model:

Test statistic	74.145
Degrees of freedom	62
P-value (Chi-square)	0.139

```
res1 = lavaan::inspect(fit1, what="std.all")
Xout = cbind(res1$lambda, diag(res1$theta))
colnames(Xout) = c("lambda_D", "lambda_A", "lambda_S", "diag(ThetaDelta)")
print(Xout)
```

	lambda_D	lambda_A	lambda_S	diag(ThetaDelta)
DASS_Q3	0.7461602	0.0000000	0.0000000	0.4432450
DASS_Q10	0.8021866	0.0000000	0.0000000	0.3564967
DASS_Q13	0.7972039	0.0000000	0.0000000	0.3644660
DASS_Q16	0.7692375	0.0000000	0.0000000	0.4082737
DASS_Q17	0.8035117	0.0000000	0.0000000	0.3543689
DASS_Q21	0.8220667	0.0000000	0.0000000	0.3242064
DASS_Q9	0.0000000	0.7245548	0.0000000	0.4750203
DASS_Q15	0.0000000	0.8093635	0.0000000	0.3449307
DASS_Q20	0.0000000	0.7491744	0.0000000	0.4387377
DASS_Q6	0.0000000	0.0000000	0.6754502	0.5437670
DASS_Q8	0.0000000	0.0000000	0.7426341	0.4484946
DASS_Q11	0.0000000	0.0000000	0.6823439	0.5344068
DASS_Q12	0.0000000	0.0000000	0.7112991	0.4940535

```
res1$psi
```

	D	A	S
D	1.000		
A	0.731	1.000	
S	0.738	0.924	1.000

```
fitMeasures(fit1, fit.measures = c("RMSEA", "CFI", "chisq", "df", "npar"))
```

	rmsea	cfi	chisq	df	npar
	0.017	0.999	74.145	62.000	55.000

Globalmente il modello adattato evidenzia buoni indici di CFI e RMSEA. La struttura fattoriale delle scale D e A è ben formata dagli item a disposizione, con coefficienti fattoriali di magnitudine sufficientemente elevata. La scala S, invece, presenta coefficienti di magnitudine lievemente inferiore alle scale precedenti (es.: DASS_Q6, DASS_Q11). Le varianze d'errore per ciascun item sono contenute ad eccezione degli item precedentemente segnalati. La matrice $\Phi_{3 \times 3}$ evidenzia che le variabili latenti sono correlate tra loro.

- In alternativa al modello adattato al punto 2, si definisca e si adatti un ulteriore modello CFA a due soli fattori latenti. Il primo fattore è relativo alla scala D mentre il secondo AS è definito usando gli item delle

scale A e S congiuntamente.

```
#ricodifichiamo ciascuna variabile osservata in termini di 'ordered factor'
mD = paste0("D =~ ", paste(paste0("DASS_Q",c(3,10,13,16,17,21)),collapse=" + "))
mAS = paste0("AS =~ ", paste(paste0("DASS_Q",c(9,15,20,6,8,11,12)),collapse=" + "))

model2 = paste(mD,mAS,sep = " \n ")
fit2 = lavaan::cfa(model = model2,data = dass21[,jld],ordered = jld,estimator="DWLS")
```

5. Si valuti se il modello a due fattori latenti sia da preferire rispetto al modello a tre fattori latenti.

```
rbind(fitMeasures(fit1,fit.measures = c("RMSEA","CFI","chisq","df","npar")),
      fitMeasures(fit2,fit.measures = c("RMSEA","CFI","chisq","df","npar")))

      rmsea      cfi      chisq df npar
[1,] 0.01728042 0.9993301 74.14519 62 55
[2,] 0.02041860 0.9990346 81.50394 64 53
```

Entrambi i modelli adattati presentano ottimi indici di adattamento. Tuttavia la soluzione a due fattori latenti presenta meno parametri da stimare ed è quindi più parsimoniosa della precedente. Sulla base di tale criterio, si preferisce la scelta di quest'ultima soluzione fattoriale.

6. Si valuti il livello di invarianza globale raggiunto dal modello CFA scelto al punto precedente rispetto alla variabile categoriale **Gender**.

```
fit2b = lavaan::cfa(model = model2,data = dass21[,c(jld,"Gender")],ordered = jld,
                    estimator="DWLS",group.equal=c("loadings"),group="Gender")
anova(fit2,fit2b)

Chi-Squared Difference Test

      Df AIC BIC      Chisq Chisq diff Df diff Pr(>Chisq)
fit2   64      188.255      81.504
fit2b 139      188.255      106.75      75 0.009397 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

L'invarianza debole non è stabilita rispetto ai dati a disposizione. Il test del χ^2 consente di rigettare l'ipotesi nulla dell'equivalenza dei due modelli CFA.

7. Si valuti mediante indici di modifica se il modello scelto al punto 5 possa essere migliorato. Successivamente, si valuti nuovamente l'invarianza debole globale.

```
X = modificationIndices(object = fit2b,free.remove = FALSE,sort. = TRUE)
print(X[1:10,])

      lhs op      rhs block group level      mi      epc sepc.lv sepc.all
305    AS =~ DASS_Q16      2      2      1 27.586 -0.247 -0.182 -0.182
168 DASS_Q16 ~*~ DASS_Q16      2      2      1 27.482 -0.224 -0.224 -1.000
72  DASS_Q16 ~*~ DASS_Q16      1      1      1 27.482 0.224 0.224 1.000
214    AS =~ DASS_Q16      1      1      1 26.743 0.238 0.156 0.156
219 DASS_Q3 ~~ DASS_Q16      1      1      1 14.598 0.161 0.161 0.326
```

```

209      D   =~ DASS_Q11      1      1      1 14.525  0.169  0.120  0.120
176 DASS_Q11 =~ DASS_Q11      2      2      1 13.314 -0.202 -0.202 -1.000
80  DASS_Q11 =~ DASS_Q11      1      1      1 13.314  0.202  0.202  1.000
4    D      =~ DASS_Q16      1      1      1 13.114  0.107  0.076  0.076
350 DASS_Q17 =~ DASS_Q21      2      2      1 10.169  0.101  0.101  0.399
      sepc.nox
305   -0.182
168   -1.000
72    1.000
214   0.156
219   0.326
209   0.120
176   -1.000
80    1.000
4     0.076
350   0.399

mD_a = paste0("D =~ ", paste(paste0("DASS_Q",c(3,10,13,16,17,21)),collapse=" + "))
mAS_a = paste0("AS =~ ", paste(paste0("DASS_Q",c(9,15,20,6,8,11,12,16)),collapse=" + "))
model2b_a = paste(mD_a,mAS_a,sep = " \n ")

fit2b_a = lavaan::cfa(model = model2b_a,data = dass21[,c(jjd,"Gender")],ordered = jjd,
                      estimator="DWLS",group="Gender",group.equal=c("loadings"))

anova(fit2b,fit2b_a)

Chi-Squared Difference Test

      Df AIC BIC  Chisq Chisq diff Df diff Pr(>Chisq)
fit2b_a 138      188.25
fit2b    139      188.26 0.00082218      1      0.9771

```

Gli indici di modifica suggeriscono che il legame `AS =~ DASS_Q16` possa essere aggiunto per entrambi i gruppi. Ciò comporta la definizione di un modello con un cross-loading (da valutare, in tal caso, la similarità semantica dei due item). Quest'ultimo modello, rispetto al precedente, permette di stabilire l'invarianza debole globale per entrambi i gruppi.

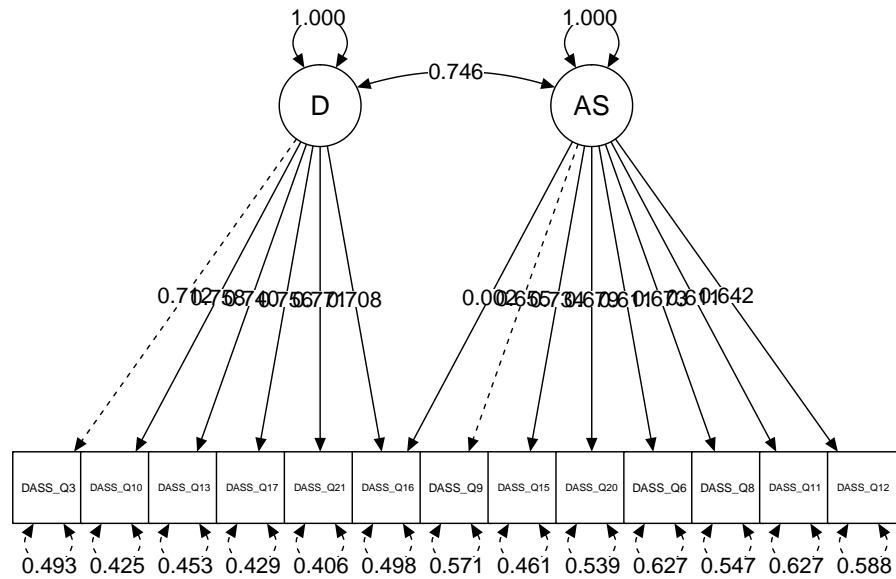
8. Si rappresenti graficamente il modello finale scelto al punto precedente.

```

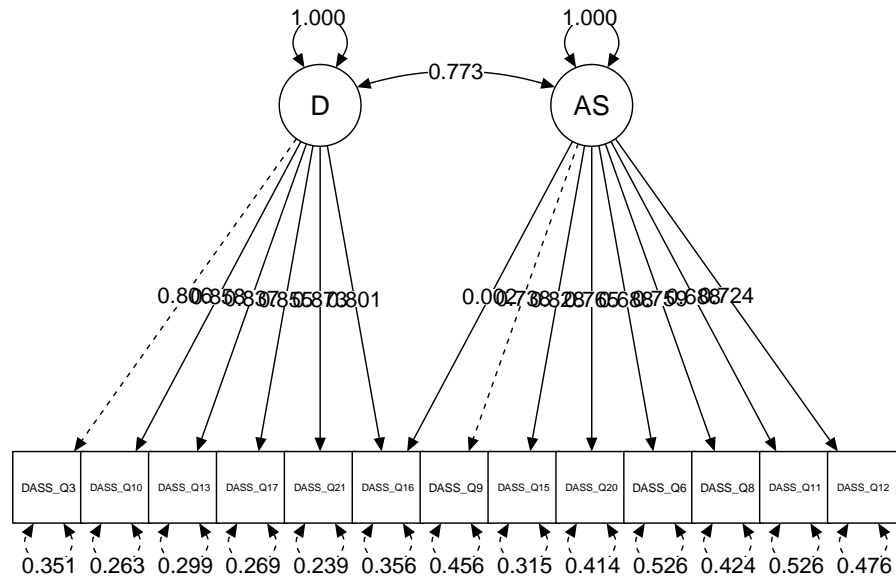
semPlot::semPaths(object = fit2b_a, whatLabels = "std.all",edge.label.cex = 0.95,
                  edge.color = "black",sizeMan = 7,sizeLat=8, style = "openmx",
                  nDigits = 3, intercepts = FALSE,thresholds = FALSE)

```

1



2



9. Si estraggano i coefficienti stimati (standardizzati) della matrice Λ per entrambi i gruppi. Si commenti perché essi differiscano tra i due gruppi considerato che il modello adattato è invariante in senso debole.

```
Lambda_est_M = lavInspect(fit2b_a,"std")$M$lambda
Lambda_est_F = lavInspect(fit2b_a,"std")$F$lambda

X = data.frame(cbind("|",round(Lambda_est_M,3),"|",round(Lambda_est_F,3)));
colnames(X) = c("", "D_M", "AS_M", "", "D_F", "AS_F")
print(X)
```

	D_M	AS_M	D_F	AS_F
DASS_Q3	0.712	0	0.806	0
DASS_Q10	0.758	0	0.858	0
DASS_Q13	0.74	0	0.837	0
DASS_Q16	0.708	0.002	0.801	0.002

DASS_Q17		0.756	0		0.855	0
DASS_Q21		0.771	0		0.873	0
DASS_Q9		0	0.655		0	0.738
DASS_Q15		0	0.734		0	0.828
DASS_Q20		0	0.679		0	0.765
DASS_Q6		0	0.611		0	0.688
DASS_Q8		0	0.673		0	0.759
DASS_Q11		0	0.611		0	0.688
DASS_Q12		0	0.642		0	0.724

Le matrici dei coefficienti fattoriali standardizzati differiscono per entrambi i gruppi anche se il modello è invariante in senso debole. Questo poiché il vincolo imposto nel modello, ossia $\mathbf{A}_M = \mathbf{A}_F$, si riferisce ai coefficienti non standardizzati, vale a dire quelli ottenuti direttamente dal processo di stima. I coefficienti standardizzati invece, $\hat{\mathbf{A}}_M$ e $\hat{\mathbf{A}}_F$, sono ottenuti dividendo i coefficienti grezzi stimati rispetto alle varianze $\text{diag}(\hat{\Phi})$ e $\text{diag}(\hat{\Sigma})$ del modello adattato. Difatti, possiamo notare che le stime non standardizzate sono le medesime per entrambi i gruppi:

```
Lambda_est_M = lavInspect(fit2b_a,"est")$M$lambda
Lambda_est_F = lavInspect(fit2b_a,"est")$F$lambda

X = data.frame(cbind("|",round(Lambda_est_M,3),"|",round(Lambda_est_F,3)));
colnames(X) = c("", "D_M", "AS_M", "", "D_F", "AS_F")
print(X)
```

		D_M	AS_M		D_F	AS_F
DASS_Q3		1	0		1	0
DASS_Q10		1.065	0		1.065	0
DASS_Q13		1.039	0		1.039	0
DASS_Q16		0.994	0.003		0.994	0.003
DASS_Q17		1.061	0		1.061	0
DASS_Q21		1.083	0		1.083	0
DASS_Q9		0	1		0	1
DASS_Q15		0	1.122		0	1.122
DASS_Q20		0	1.037		0	1.037
DASS_Q6		0	0.933		0	0.933
DASS_Q8		0	1.028		0	1.028
DASS_Q11		0	0.933		0	0.933
DASS_Q12		0	0.981		0	0.981

10. Si calcoli l'attendibilità del test scelto al punto 7 e si commenti il risultato ottenuto.

```
semTools::reliability(object = fit2b_a,what = "omega",return.total = TRUE)

$M
      D      AS    total
omega 0.8466911 0.7759684 0.8933231

$F
      D      AS    total
omega 0.9085776 0.8228298 0.9317209
```

L'attendibilità stimata mediante indice ω mostra un'alta attendibilità per entrambe le sottoscale del test. Complessivamente, per entrambi i gruppi, l'attendibilità è elevata e ciò indica che il test quantifica accu-

ratamente i costrutti D e AS.

SOLUZIONE