

Tecniche di ricerca psicologica e analisi dei dati

A.A. 2022/23

PROF.SSA LISA DI BLAS

diblali@units.it

Il semestre: 3 variabili

Contenuti principali:

- Metodologia della ricerca: ruoli e relazioni tra variabili
- ANOVA between con 2 VI o disegni fattoriali
- ANOVA per disegni fattoriali misti
- Analisi della regressione multipla
- Tecniche di riduzione dei dati

Tipi di variabili e di relazioni tra variabili

Tipi di variabili e di relazioni tra variabili

2 variabili

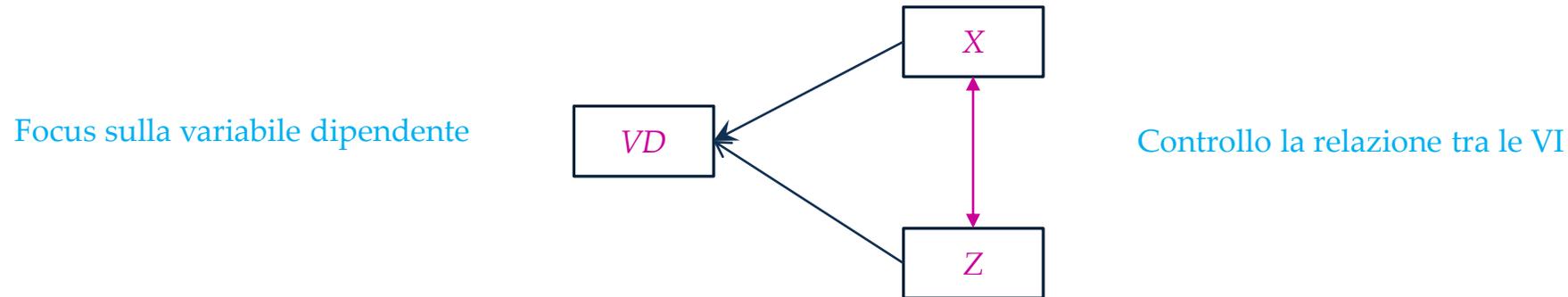
- co-occorrenza
- dipendenza
- causalità
- indicazione

3 variabili

- effetto additivo dato dall'impatto unico di ogni stimatore (parzializzazione o depurazione o controllo della relazione tra variabili indipendenti)
- interazione o moderazione
- mediazione
- relazioni depurata di tipo spurio e di tipo soppresso (ma anche relazione rovesciata e deflazionata)

3 variabili: quale il ruolo di ogni variabile?

- se vi sono più VI, si osserva un **effetto additivo** quando ciascuna variabile contribuisce a spiegare parte della variabilità di Y, al di là della relazione tra le VI, senza che vi sia interazione



Esempi:

I livelli attentivi variano in funzione delle ore di sonno e dell'età

Motivazione all'apprendimento e intelligenza contribuiscono al successo scolastico

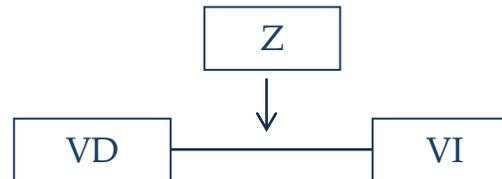
Senso di solitudine e bassa auto-stima spiegano differenze individuali nella condizione depressiva

3 variabili : quale il ruolo di ogni variabile?

Tipi di variabili: intervenire in una relazione per **controllare** e **stratificare**

- se vi sono più VI, però, potremmo chiederci se esse interagiscono, vale a dire **se l'effetto di una VI (es. X) sulla VD varia a seconda dei livelli dell'altra VI (es. Z)**. Un **effetto d'interazione** indica che i livelli attesi di VD variano in funzione delle possibili combinazioni delle VI.
- se le VI interagiscono, allora una VI (Z) funge da **variabile moderatrice** e la relazione tra X e Y si dice moderata da Z: L'intensità della relazione tra Y e X varia in funzione dei livelli di VMod (Z)

Focus sulla relazione tra VD (Y) e VI (X)



Esempi:

I livelli attentivi variano in funzione delle ore di sonno solo in soggetti di età inferiore ai 30 anni

Un testimonial attraente (vs neutro) genera un atteggiamento favorevole verso un prodotto solo se la pubblicità presenta argomentazioni forti (vs deboli)

Tanti modi per esprimere lo stesso effetto

I livelli attentivi variano in funzione delle ore di sonno solo in soggetti di età inferiore ai 30 anni

Le capacità attentive risentono delle ore di sonno quando si è giovani.

Le ore di sonno possono favorire o indebolire i livelli attentivi nei giovani

Un testimonial attraente (vs neutro) genera un atteggiamento favorevole verso un prodotto solo se la pubblicità presenta argomentazioni forti (vs deboli)

Quando un pubblicità utilizza argomentazioni forti, allora non è necessario un testimonial attraente per favorire un atteggiamento positivo verso un prodotto.

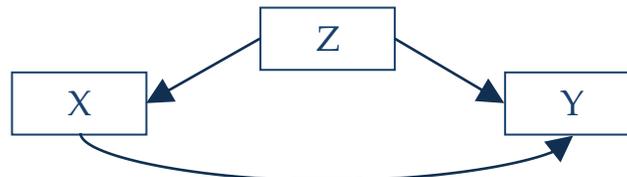
La condizione sperimentale abbassa lo humor aggressivo soprattutto negli uomini

Il trattamento sperimentale agisce con maggiore efficacia sullo humor aggressivo negli uomini.

3 variabili: quale il ruolo di ogni variabile?

Tipi di variabili: intervenire in una relazione per depurare e spiegare:

- **variabile interveniente e relazione spuria**: la relazione tra VD e VI (Y e X) è indiretta: X si associa a Y per la **variabile interveniente** (Z) che si associa a entrambe e precede entrambe logicamente



Esempio

Il numero di rondini in cielo e il numero di matrimoni è correlato, ma solo se non si controlla la stagione (primavera e autunno sono le stagioni di migrazione delle rondini e quelle privilegiate per sposarsi)

3 variabili: quale il ruolo di ogni variabile?

Tipi di variabili: intervenire in una relazione per **depurare e spiegare**:

- variabile mediatrice e relazione mediata: la relazione tra VD e VI (Y e X) è indiretta: X agisce su Y (almeno in parte) attraverso la **variabile mediatrice (Z)**



Esempio:

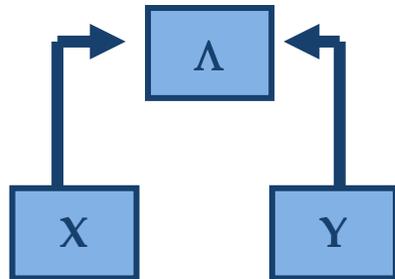
L'ambiente sociale incide su output comportamentali bimbo indirettamente, attraverso la mediazione dell'ambiente familiare

Le basi genetiche dell'aggressività ereditata dai genitori agiscono indirettamente sui bambini attraverso l'ambiente familiare aggressivo sviluppato dal genitore stesso

3 variabili: quale il ruolo di ogni variabile?

Costrutto emerge dagli indicatori comportamentali:

Relazione di indicazione

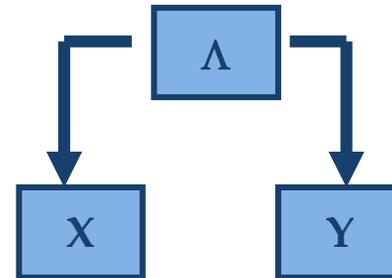


Indicatori formativi

Ansia e fragilità emotive sono indicatori di Instabilità emotiva.

Costrutto latente causa le variazioni negli indicatori:

Relazione di dipendenza



Indicatori riflettivi

Ansia e fragilità emotive dipendono dall'Instabilità emotiva.

ANOVA

per disegni fattoriali between e misti

3 variabili: ANOVA

- VD = quantitativa
- (almeno) 2 VI = qualitative (dicotomiche) di cui 1 funge da V Moderatrice (dicotomica)
- VI e VM possono essere
 - tra i soggetti (gruppi/condizioni) (BTW)
 - variabili sperimentali
 - non sperimentali
 - entro i soggetti (gruppi/condizioni)(WTH)
 - variabili sperimentali
 - variabile temporale o misure ripetute nel tempo
- ANOVA BTW \rightarrow 1 VI_{BTW} e 1 VM_{BTW}
- ANOVA MISTA \rightarrow 1 VI_{WTH} e 1 VM_{BTW}

Disegni (sperimentali) tra o BTW

- sono detti disegni fattoriali
- le VI sono chiamate “fattori” e hanno almeno 2 livelli
- si verificano effetti principali ed effetti d’interazione
- i Ss sono assegnati in modo casuale alle possibili combinazioni delle VI
- notazione: si indicano le VI separate da un x, riferendone i livelli (es., 2 x 2 significa che ci sono due VI, X e Z, ciascuna delle quali ha 2 livelli)
- il numero di fattori e rispettivi livelli deve essere contenuto in ragione sia di N sia di elementi di tipo statistico

I disegni bilanciati

se n è costante attraverso le celle, il disegno è bilanciato

→ indipendenza o ortogonalità tra VI tenute,

costanti attraverso le condizioni grazie al disegno metodologico

5 m lavoratori	5 f lavoratrici
5 m non lavoratori	5 f non lavoratrici

Disegni sperimentali between con 2 VI

Un esempio

In un disegno sperimentale, si vuole verificare se l'immagine corporea (VD) varia in funzione di

Fattore X: la VD varia in funzione di X, indipendentemente o al di là delle differenze rispetto al fattore Z? ?

Fattore Z: la VD varia in funzione di Z, indipendentemente o al di là delle differenze rispetto al fattore X? ?

NB Si deve stabilire quale VI fungerà da VI e quale da Vmoderatrice

Possibile interazione, vale a dire i livelli di Y variano a seconda di come si combinano Fattore X e Fattore Z (es. Ginnastica favorisce benessere psicofisico **se** si combina con una dieta da zuccheri? Il trattamento sperimentale sullo humor aggressivo ha effetti simili o no tra uomini e donne?)

Maggiore punteggio Post test, maggiore la soddisfazione corporea

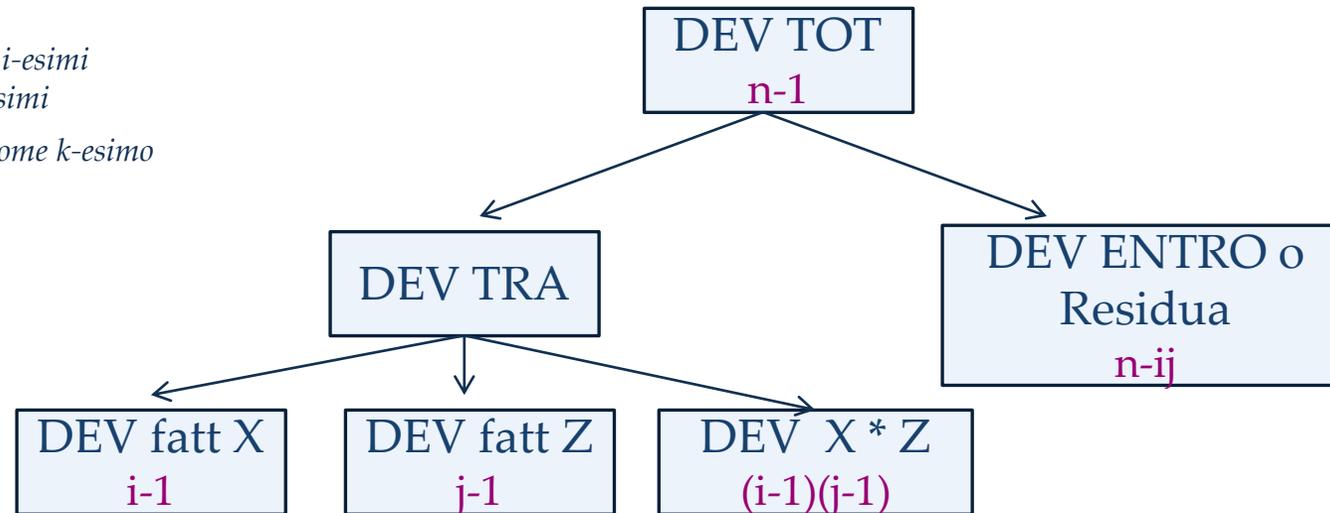
Disegni sperimentali between con 2 VI

Si può dimostrare che $DEV\ TOT = DEV\ TRA + DEV\ ENTRO$

e si scompone come segue

dove

- la variabile X ha livelli i-esimi
- la variabile Z livelli j-esimi
- l'individuo è indicato come k-esimo



La variabilità totale dipende dalla variabilità del fattore X, del fattore Z e della loro combinazione X * Z (interazione)

$$y_{ijk} = M_{GEN} + (M_{gruppo_i} - M_{GEN}) + (M_{gruppo_j} - M_{GEN}) + \\ + [M_{gruppo_ij} - M_{GEN} - (x_i + z_j)] + (y_{ijk} - M_{gruppo_ij})$$

ovvero in termini di scomposizione della devianza totale (SQ)

$$SQ_y(y_{ijk} - M_{GEN}) = SQ_X(M_{gruppo_i} - M_{GEN}) + SQ_Z(M_{gruppo_j} - M_{GEN}) + \\ + SQ_{X*Z}[SQ_{Between} - (SQ_X + SQ_Z)] + SQ_{Residua}(y_{ijk} - M_{gruppo_ij})$$

dove $SQ_{Between} = (M_{gruppo_ij} - M_{GEN}) = SQ_X + SQ_Z + SQ_{X*Z}$

Disegni sperimentali between con 2 VI

- effetto principale del fattore X (*y* varia in funzione di X al di là di Z?)
- effetto principale del fattore Z (*y* varia in funzione di Z al di là di X?)
- effetto d'interazione $X * Z$ vale a dire, *y* varia in funzione di quella combinazione particolare *ij* che non dipende né da elementi residuali né dagli effetti principali né dall'effetto generale o media generale?

File excel + file Jamovi

ANOVA 2 x 2 per verificare la scomposizione della devianza totale

Disegni sperimentali between con 2 VI

$$R^2 = SQ_{BTW} / SQ_{TOT}$$

$$\text{Test F} = (SQ_{BTW}/gl_{BTW}) / (SQ_{TOT}/gl_{TOT})$$

dove $gl_{BTW} = (\text{livelli X} * \text{livelli Z}) - 1$

$$gl_{TOT} = N - (\text{livelli X} * \text{livelli Z})$$

$$p\eta^2_X = SQ_X / (SQ_X + SQ_{ERRORE})$$

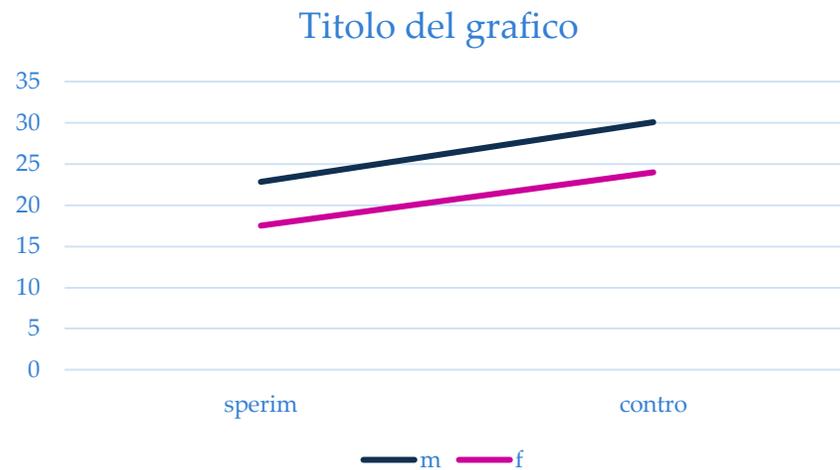
$$\text{test F} = (SQ_X/gl_X) / (SQ_{RES}/gl_{RES})$$

$$p\eta^2_{X*Y} = SQ_{X*Y} / (SQ_{X*Y} + SQ_{ERRORE})$$

$$\text{test F} = (SQ_{XY}/gl_{XY}) / (SQ_{RES}/gl_{RES})$$

Esempio file Jamovi: VD Umore aggressivo VI condizione sperimentale Vmod Genere

La condizione sperimentale abbassa lo humor aggressivo allo stesso modo tra uomini e donne?



Descriptives

Descriptives

	Genere	Sperimentale	Aggressive_Humor
N	m	speriment	21
		controllo	21
	f	speriment	21
		controllo	21
Mean	m	speriment	22.8
		controllo	30.1
	f	speriment	17.5
		controllo	24.0
Standard deviation	m	speriment	5.82
		controllo	6.51
	f	speriment	5.44
		controllo	4.13

I risultati suggeriscono che mediam lo humor aggress è minore nella condz sperimentale vs controllo ed è mediam minore nelle donne rispetto agli uomini; pare nessun effetto di interazione; è cosi?

Esempio file Jamovi

La condizione sperimentale abbassa lo humor aggressivo allo stesso modo tra uomini e donne?

Estimated Marginal Means

Genere

Estimated Marginal Means - Genere

Genere	Mean	SE	95% Confidence Interval	
			Lower	Upper
0	26.5	0.855	24.8	28.2
1	20.7	0.855	19.0	22.4

Sperimentale

Estimated Marginal Means - Sperimentale

Sperimentale	Mean	SE	95% Confidence Interval	
			Lower	Upper
0	20.1	0.855	18.4	21.8
1	27.0	0.855	25.3	28.7

ANOVA - Aggressive_Humor

	Sum of Squares	df	Mean Square	F	p	η^2	η^2p
Overall model	1689.18	3	563.06	18.331	< .001		
Genere	691.44	1	691.44	22.511	< .001	0.167	0.220
Sperimentale	994.30	1	994.30	32.371	< .001	0.240	0.288
Genere * Sperimentale	3.44	1	3.44	0.112	0.739	0.001	0.001
Residuals	2457.24	80	30.72				

L'effetto del trattamento sperimentale sullo humor non varia tra uomini e donne; genere e trattamento hanno un effetto additivo sullo humor aggressivo

Un altro esempio:
 VD = benessere psicofisico
 VI = Attività fisica
 Vmod= Dieta da zuccheri

*L'attività fisica ha un effetto sul benessere soprattutto se combinata con una dieta da zuccheri?
 Lo studio è sperimentale*

Descriptives

Descriptives

	Ginnastica	Dieta	Post-test
N	si	si	10
		no	10
	no	si	10
		no	10
Mean	si	si	54.3
		no	53.3
	no	si	54.0
		no	46.5
Standard deviation	si	si	4.62
		no	3.53
	no	si	5.60
		no	4.43



La dieta sembra avere un effetto protettivo rispetto alla mancanza di ginnastica, è così?

ANOVA - Post-test

	Sum of Squares	df	Mean Square	F	p	η^2	η^2p
Overall model	412	3	137.4	6.49	0.001		
Dieta	181	1	180.6	8.53	0.006	0.154	0.191
Ginnastica	126	1	126.0	5.95	0.020	0.107	0.142
Dieta * Ginnastica	106	1	105.6	4.99	0.032	0.090	0.122
Residuals	763	36	21.2				

Ginnastica

Estimated Marginal Means - Ginnastica

Ginnastica	Mean	SE	95% Confidence Interval	
			Lower	Upper
si	53.8	1.03	51.7	55.9
no	50.2	1.03	48.2	52.3

Estimated Marginal Means - Dieta

Dieta	Mean	SE	95% Confidence Interval	
			Lower	Upper
si	54.1	1.03	52.1	56.2
no	49.9	1.03	47.8	52.0

I risultati confermano quanto suggerito dal grafico: dieta sembra avere un effetto protettivo rispetto alla mancanza di ginnastica. Gli effetti principali sono significativi, ma i risultati si leggono in relazione all'interazione

Post Hoc Comparisons - Dieta * Ginnastica

Comparison					Mean Difference	SE	df	t	p _{Tukey}
Dieta	Ginnastica	Dieta	Ginnastica						
si	si	-	si	no	0.300	2.06	36.0	0.146	0.999
		-	no	si	1.000	2.06	36.0	0.486	0.962
		-	no	no	7.800	2.06	36.0	3.789	0.003
	no	-	no	si	0.700	2.06	36.0	0.340	0.986
		-	no	no	7.500	2.06	36.0	3.644	0.004
		-	no	no	6.800	2.06	36.0	3.303	0.011

I risultati post-hoc confermano che il gruppo di partecipanti Assegnato alla condizione No Dieta e No Ginnastica riferisce livelli inferiori di Benessere rispetto a tutte le altre possibili combinazioni

Note. Comparisons are based on estimated marginal means

Il test statistico HSD (honestly significant difference) di Tukey è attendibile quando il numero di confronti è < 5 e non risente del numero di confronti operati

Disegni quasi-sperimentali between

o disegni ex post facto o con gruppi non equivalenti

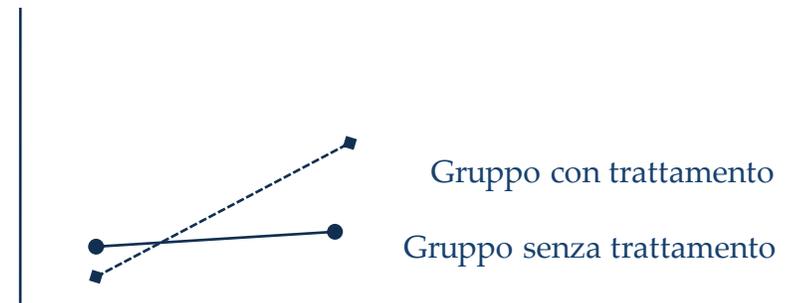
- non è possibile manipolare la VI, ma definirne almeno due livelli si
- e assegnare i Ss ai gruppi, seppure non in modo casuale, poiché formano entità predefinite (es. studenti di 2 scuole, poste nella stessa area urbana) o anche naturali (es., m/f, dove genere funge da VI)
- attuare procedure specifiche di tipo sperimentale, con disegni pre-test/trattamento/post-test, per verificare le ipotesi (es. gli studenti di una scuola sono sottoposti a trattamento, quelli dell'altra scuola no)
- includere controlli sulle minacce alla validità (interna, esterna, ma anche statistica e di costrutto)
- è possibile verificare l'equivalenza statistica tra i gruppi
- tuttavia anche quando verificata non si può escludere del tutto che non la VI ma altre variabili *confounded* abbiano agito sugli esiti del post-test

Disegni quasi-sperimentali between

- è possibile verificare l'equivalenza statistica tra i gruppi
- tuttavia anche quando verificata non si può escludere del tutto che non la VI ma altre variabili *confounded* abbiano agito sugli esiti del post-test
- pertanto si verifica l'efficacia della VI trattamento in termini differenze pre-post test osservate tra i gruppi studiati

	Pre-test	Trattamento	Post-test	Differenza
Gruppo A (Nr)	si	si	si	pre-post
Gruppo B (Nr)	si	no	si	pre-post

il *cross-over effect* garantisce
più di altri risultati l'efficacia
del trattamento → ANOVA con disegno misto



Disegni fattoriali misti con 1 VI within e 1 VI (moderatrice) between

In questi disegni, i diversi livelli di una variabile sono somministrati agli stessi partecipanti (Fattore within), mentre i livelli dell'altra variabile no e distinguono i partecipanti in gruppi (Fattore between).

Fattore within può essere

- sperimentale (trattamento e controllo somministrati agli stessi partecipanti)
- non sperimentale, con somministrazione ripetuta nel tempo della stessa variabile (VI = tempo)

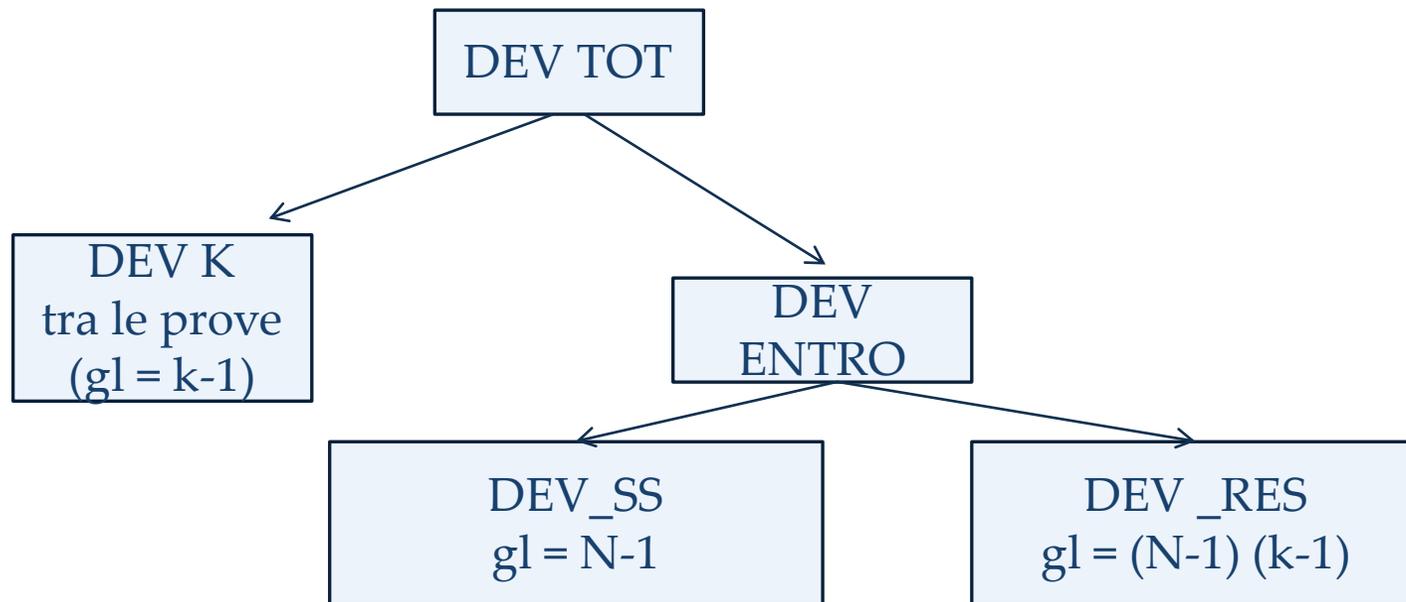
Fattore between può essere

- sperimentale (trattamento e controllo a gruppi distinti di partecipanti, assegnati casualmente ai diversi livelli)
- non sperimentale (genere, gruppi pre-definiti)

I livelli di autostima variano da un anno all'altro più rapidamente nelle pre-adolescenti che nei pre-adolescenti.

Il riconoscimento facciale corretto è mediamente più rapido nella condizione luminosità media piuttosto che elevata.

Disegni sperimentali within con 1 VI

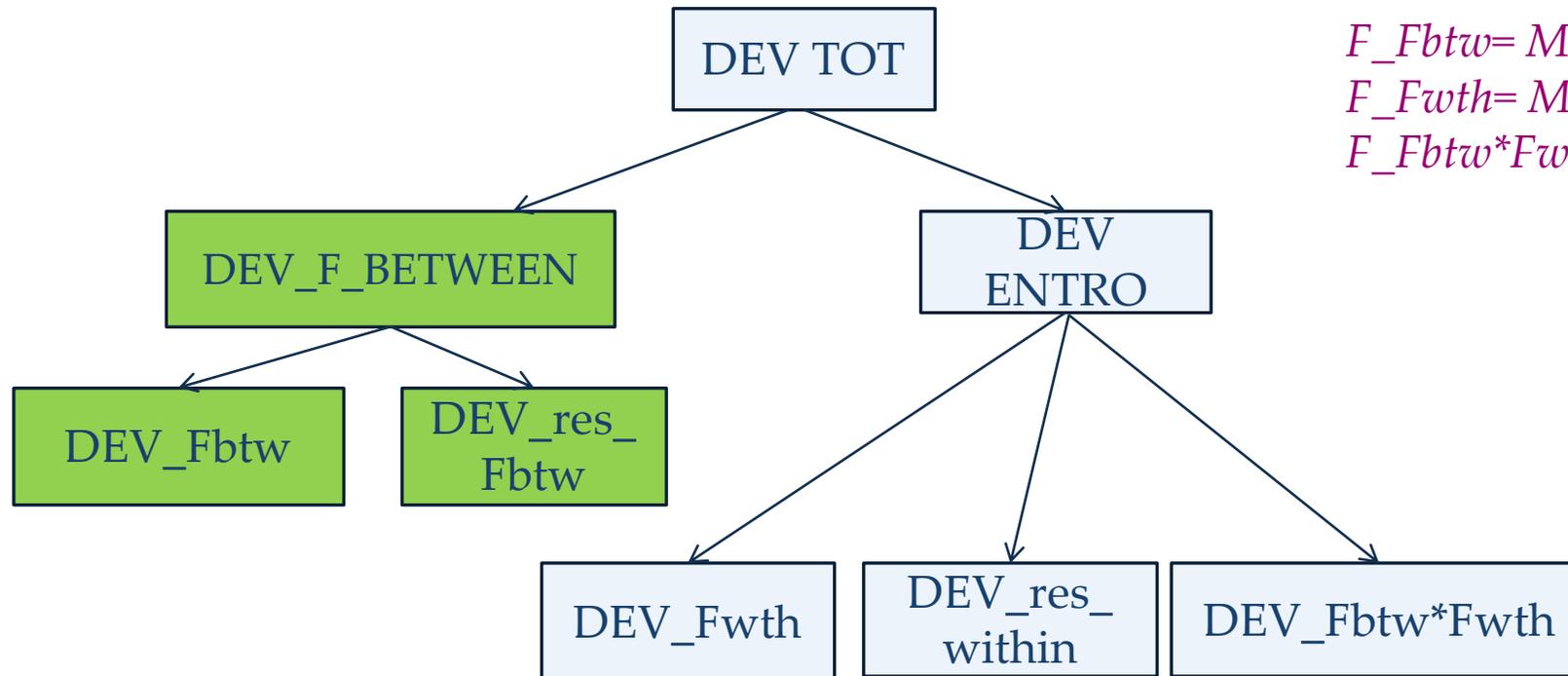


Nell'ANOVA per misure ripetute il rapporto F corrisponde a

$$F = (DEV_K/gl) / (DEV_{RES}/gl)$$

Disegni misti con 1 VI within e 1 VI between

Scomposizione della varianza

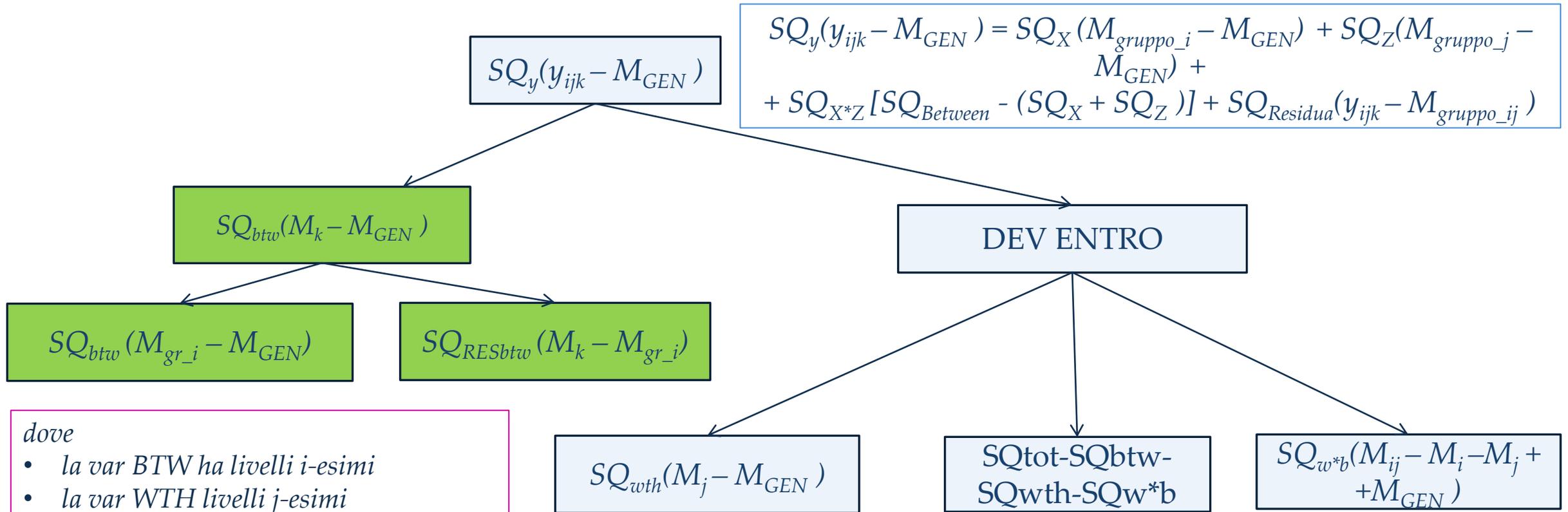


$$F_{Fbtw} = MS_{Fbtw} / MS_{res_Fbtw}$$

$$F_{Fwth} = MS_{Fwth} / MS_{res_Fwth}$$

$$F_{Fbtw} * F_{Fwth} = MS_{Fbtw} F_{Fwth} / MS_{res_within}$$

Disegni misti con 1 VI within e 1 VI between



- dove
- la var BTW ha livelli i-esimi
 - la var WTH livelli j-esimi
 - l'individuo è indicato come k-esimo
 - $S = \sum_i \sum_j \sum_k$

Esempio: file Jamovi ANOVA MISTA

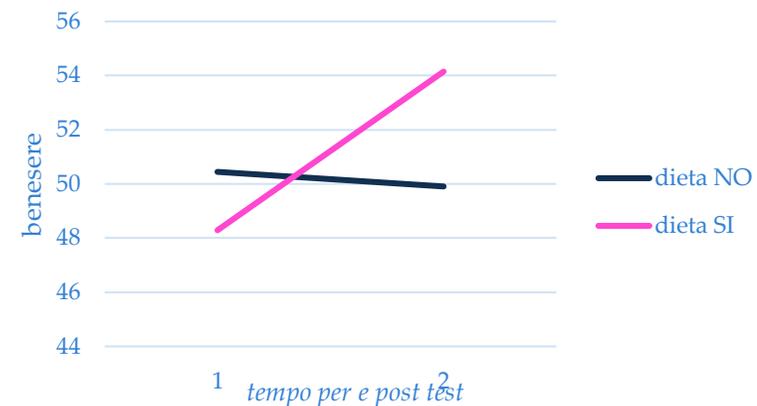
In un campione di pz ospedalizzati una dieta bilanciata ha un impatto sul loro benessere nell'arco di tre settimane? (dati fittizi)

VD = benessere psicologico

VIwth = tempo, 2 livelli (valutazione benessere pre- e post trattamento)

VMod Btw = trattamento Dieta a 2 livelli (si vs no, con assegnazione casuale ai gruppi)

	media		dieta NO	dieta SI
pre-test	49.375		50.45	48.3
post-test	52.025		49.9	54.15
a dieta	51.225			
non a dieta	50.175			



I risultati descrittivi suggeriscono un incremento medio nel tempo di benessere e un effetto però di interazione, Il benessere aumenta nei pz sottoposti a dieta bilanciata; Il fattore btw pare trascurabile. È così?

Risultati elaborati con Jamovi

Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η^2	η^2_p
benessere	140.4	1	140.45	55.2	< .001	0.059	0.592
benessere * Dieta	204.8	1	204.80	80.4	< .001	0.086	0.679
Residual	96.7	38	2.55				

Note. Type 3 Sums of Squares

è statist sig sia l'effetto principale «tempo» rispetto al benessere sia l'interazione «tempo» x dieta

Between Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η^2	η^2_p
Dieta	22.0	1	22.0	0.439	0.511	0.009	0.011
Residual	1906.7	38	50.2				

Note. Type 3 Sums of Squares

Mediamente i pz assegnati ai due gruppi «dieta» non differiscono tra loro

Post Hoc Comparisons - benessere * Dieta

Comparison					Mean Difference	SE	df	t	p _{tukey}
benessere	Dieta	benessere	Dieta						
Level 1	si	-	Level 1	no	-2.150	1.630	38.0	-1.319	0.557
		-	Level 2	si	-5.850	0.505	38.0	-11.594	< .001
		-	Level 2	no	-1.600	1.624	38.0	-0.985	0.759
	no	-	Level 2	si	-3.700	1.624	38.0	-2.279	0.121
		-	Level 2	no	0.550	0.505	38.0	1.090	0.698
		-	Level 2	no	4.250	1.618	38.0	2.627	0.057

Alcune note e assunzioni sui disegni fattoriali

- il bilanciamento di un disegno di ricerca (stesso n per ogni cella, mantenendo costante la distribuzione attraverso i livelli delle VI) fa sì che gli effetti principali e di interazione siano ortogonali o indipendenti, c'è un controllo della co-variazione tra le variabili grazie al bilanciamento dei partecipanti; se i disegni non sono bilanciati quel controllo viene meno e l'applicazione della tecnica della regressione rappresenta un rimedio (nell'ANOVA «tipo III» analitico)
- distribuzione normale della variabile osservata VD nei gruppi
 - N
 - outlier
 - Test di Kolmogorov-Smirnov e test di Shapiro-Wilk
- omoschedasticità: la varianza nei diversi gruppi o livelli della VI è comparabile statisticamente
 - Test di Levene
 - Test di sfericità (Mauchley) per disegni a misure ripetute con VI with a 3 o più livelli (la varianza delle differenze tra tutte le copie delle misure ripetute è uguale)