

Analisi delle medie post ANOVA

(cap. 4 e 8)

Marcello Gallucci

marcello.gallucci@unimib.it

AMD

Problema

- Se il test F è significativo, concludiamo che esiste una differenza fra le medie dei gruppi
- Ciò vuol dire che almeno due medie sono differenti, ma non necessariamente che tutte le medie sono differenti
- Chiameremo questo effetto *Effetto principale* (main effect)
- Ci manca di sapere dove esattamente sono queste differenze

- Analisi dei contrasti
 - Se abbiamo ipotesi specifiche da testare riguardanti gruppi di medie
- Test post-hoc
 - Se volgiamo confrontare tutte le medie con tutte le altre

Disegni Fattoriali

- Nelle maggior parte delle applicazioni di ricerca, il disegno di ricerca prevede più variabili indipendenti **incrociate**

Disegno fattoriale 3 (livello di stress) X 2 (genere)

		Genere	
		Maschi	Femmine
Variabile Sperimentale	Controllo	G1	G2
	Stress Emotivo	G3	G4
	Stress Cognitivo	G5	G6

Nei disegni fattoriali ogni **gruppo** di partecipanti rappresenta una **combinazioni di livelli** delle variabili indipendenti

- Analisi dei contrasti
 - Se abbiamo ipotesi specifiche da testare riguardanti gruppi di medie
- Test post-hoc
 - Se volgiamo confrontare tutte le medie con tutte le altre

Metodo dei contrasti
(trend analysis ed altro)

Medie

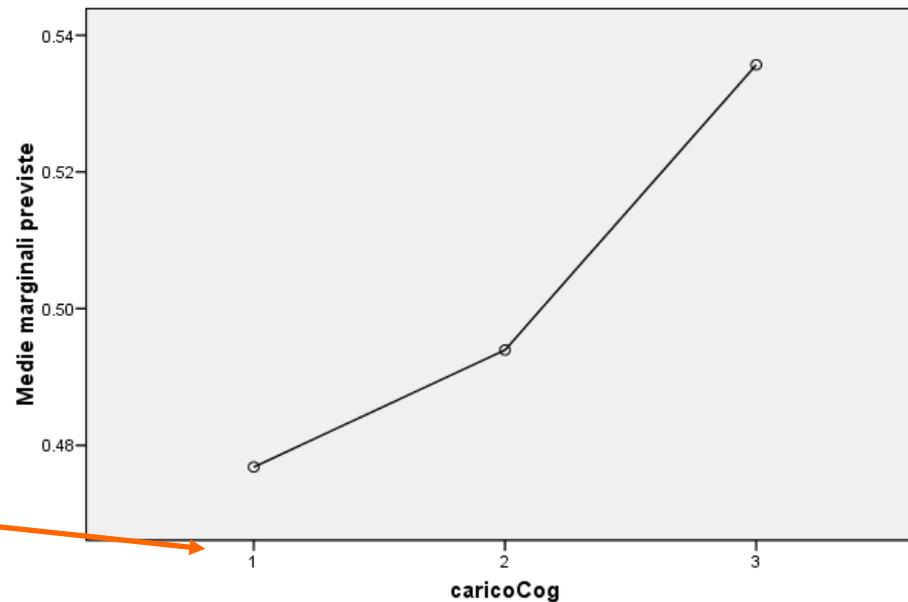
Guardo i dati: guardando il grafico delle medie, capisco cosa accade alla proporzione di errori per i diversi carichi cognitivi

Grafici di profilo

Quali medie sono diverse fra loro?

Come cambiano le medie all'aumentare del carico

Medie marginali previste di MEASURE_1



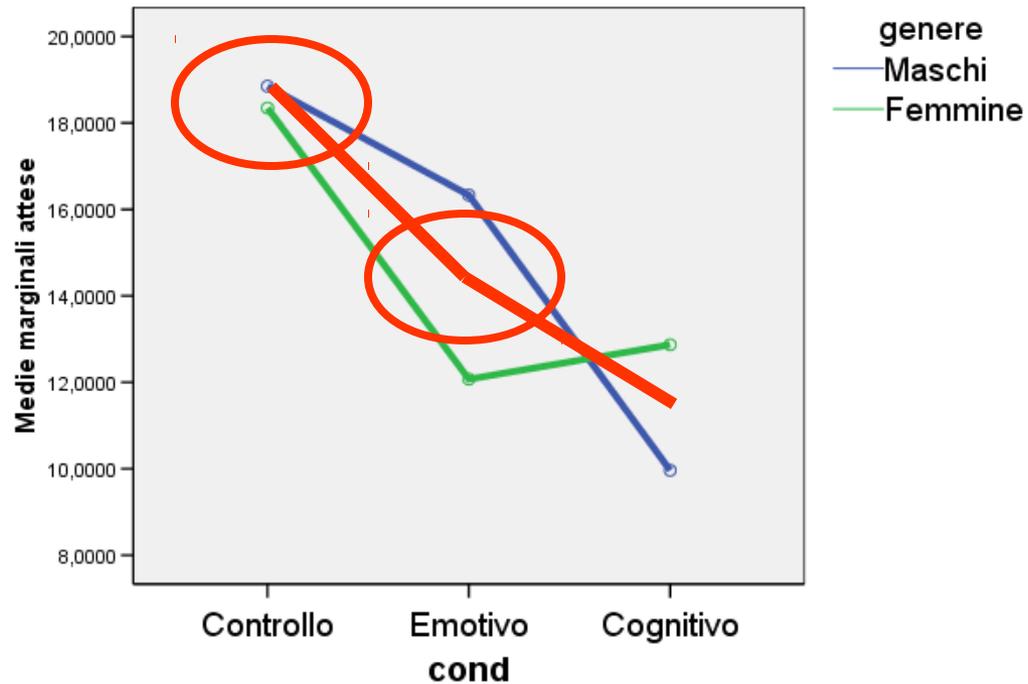
Livello di stress cognitivo

Contrasti

- Consideriamo l'effetto principale di condizione

Controllo è diverso da Emotivo?

Medie marginali attese di perform



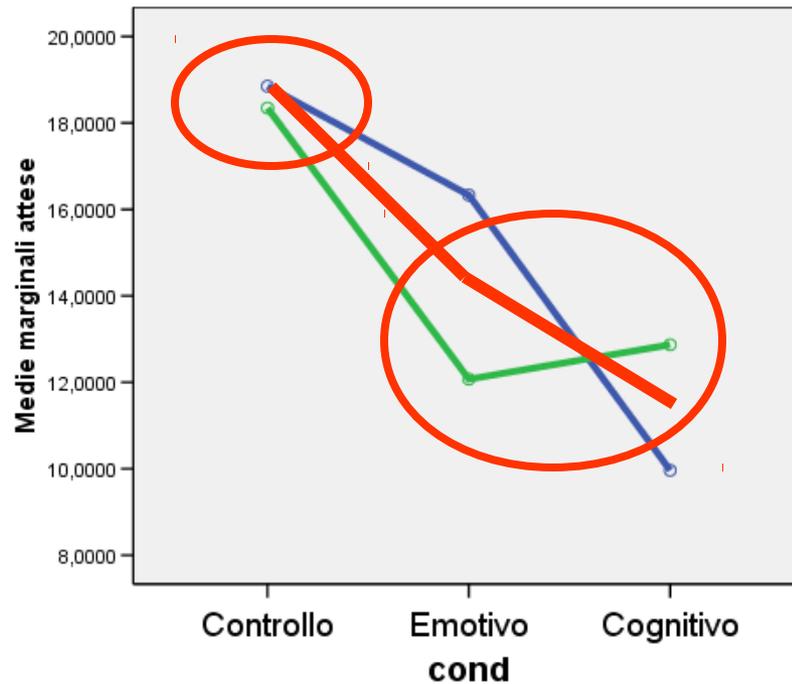
cond	stress LSMEAN
1	18.5962800
2	14.2012125
3	11.4167850

Contrasti

- Consideriamo l'effetto principale di condizione

Controllo è diverso da stress (Emotivo e Cognitivo)?

Medie marginali attese di perform



cond	stress LSMEAN
1	18.5962800
2	14.2012125
3	11.4167850

Costruire un contrasto

- Un contrasto è un set di coefficienti (pesi) che specificano un'ipotesi nulla che vogliamo rifiutare

- 1) *Controllo è diverso da Emotivo?*
- 2) *Controllo è diverso da stress (Emotivo e Cognitivo)?*

- Scriviamo tale relazione in modo tale che se le medie fossero tutte uguali (l'ipotesi nulla), la somma delle medie sarebbe zero.

$$CONT - EMOT = 0 \qquad CONT - \frac{EMOT + COGN}{2} = 0$$

I coefficienti dei contrasti sono semplicemente i numeri che moltiplicano le medie nella nostra ipotesi

	CONT	EMOT	COGN		CONT	EMOT	COGN
Cont1	1	-1	0	Cont2	1	-1/2	-1/2

Costruire un contrasto

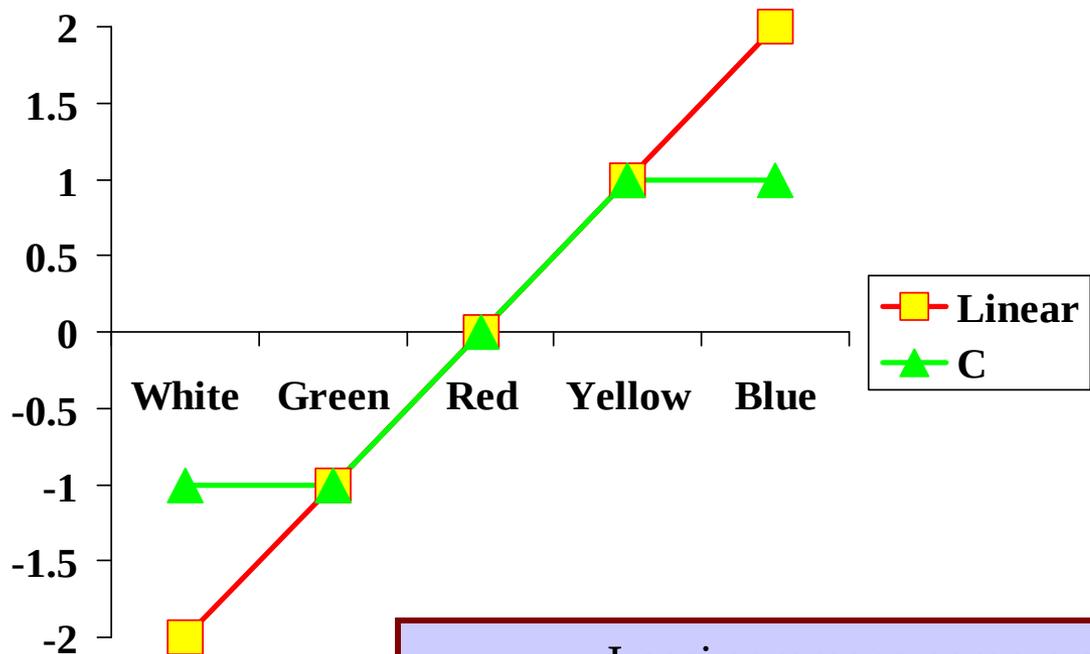
- Stimare un contrasto equivale a costruire una nuova variabile sommi le medie dei gruppi pesandoli (moltiplicandoli) per dei coefficienti
- I coefficienti hanno somma zero, e dunque testano una ipotesi nulla
- L'ipotesi nulla testata è la differenza tra le medie con peso positivo e quelle con peso negativo

1)

	CONT	EMOT	COGN		CONT	EMOT	COGN
Cont1	1	-1	0	Cont2	1	-1/2	-1/2

Contrasti in generale

In generale possiamo usare qualunque peso purchè descriva un'ipotesi nulla



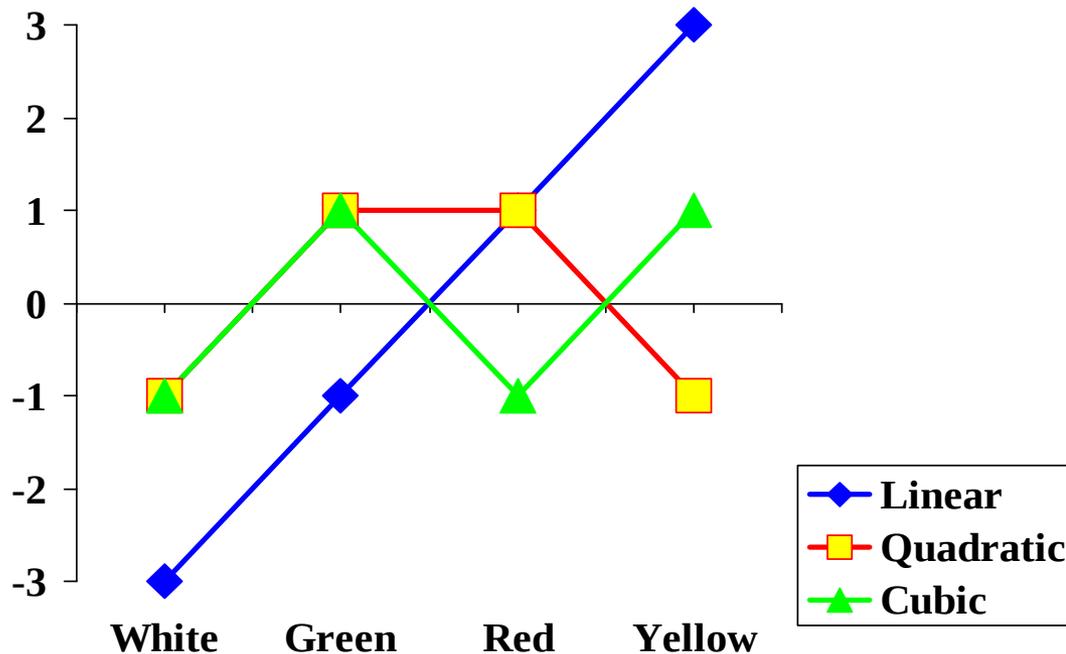
$$L = \begin{pmatrix} -3 & -1 & 0 & 1 & 3 \end{pmatrix}$$

$$C = \begin{pmatrix} -1 & -1 & 0 & 1 & 1 \end{pmatrix}$$

I pesi sommano sempre a zero

Contrasti Polinomiali

I contrasti polinomiali sono un set di pesi noto in quanto sono ortogonali fra loro e studiano il trend delle medie



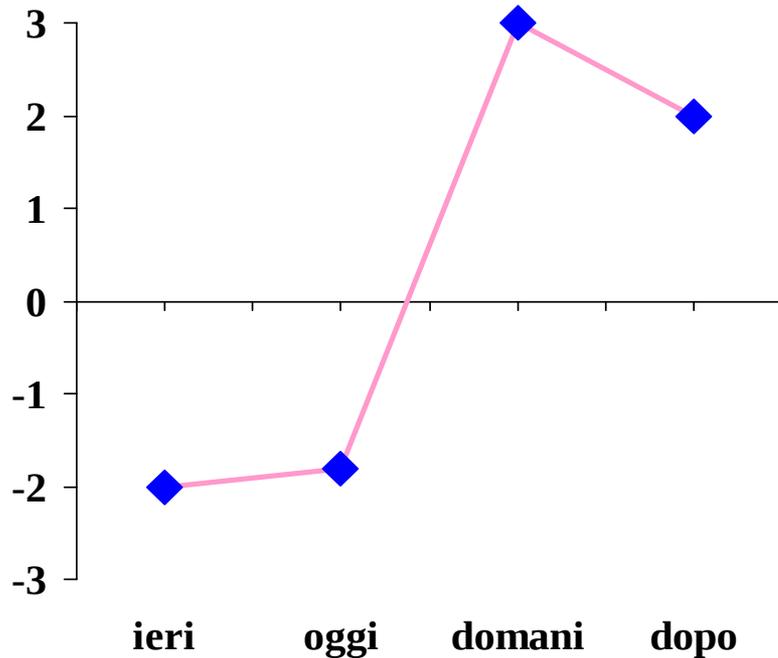
$$L = \begin{pmatrix} -3 & -1 & 1 & 3 \end{pmatrix}$$

$$Q = \begin{pmatrix} -1 & 1 & 1 & -1 \end{pmatrix}$$

$$C = \begin{pmatrix} -1 & 1 & -1 & 1 \end{pmatrix}$$

Contrasti Polinomiali

Cioè ogni polinomio (contrasto) risponde ad una domanda



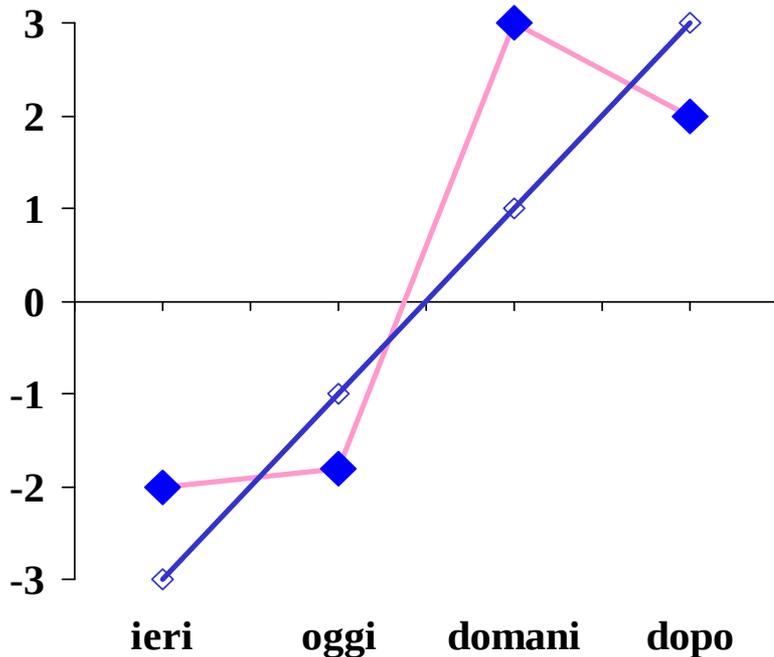
$$L = \begin{pmatrix} -3 & -1 & 1 & 3 \end{pmatrix}$$

$$Q = \begin{pmatrix} -1 & 1 & 1 & -1 \end{pmatrix}$$

$$C = \begin{pmatrix} -1 & 1 & -1 & 1 \end{pmatrix}$$

Contrasti Polinomiali

Cioè ogni polinomio (contrasto) risponde ad una domanda: c'è un trend lineare tra le medie?



Lineare

$$L = \begin{pmatrix} -3 & -1 & 1 & 3 \end{pmatrix}$$

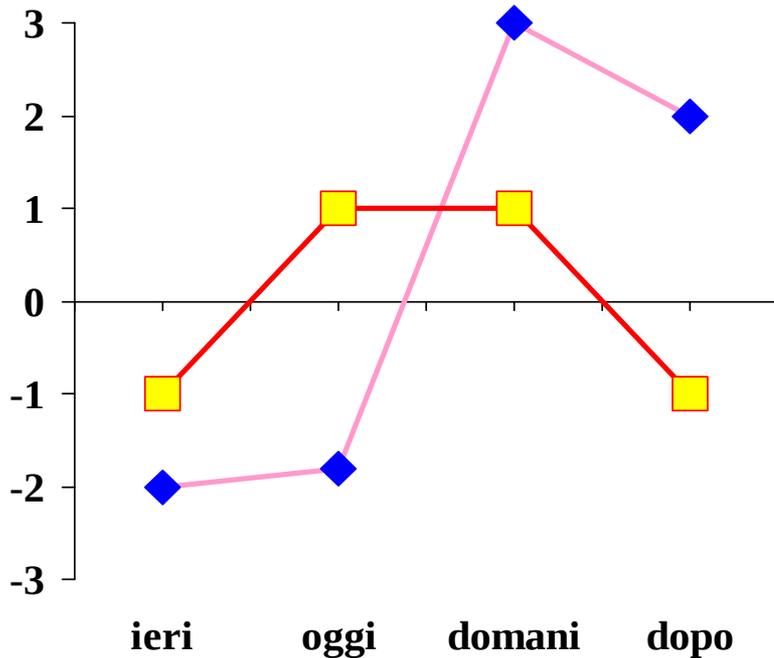
$$Q = \begin{pmatrix} -1 & 1 & 1 & -1 \end{pmatrix}$$

$$C = \begin{pmatrix} -1 & 1 & -1 & 1 \end{pmatrix}$$

C'è una tendenza nel trend a salire?

Contrasti Polinomiali

Cioè ogni polinomio (contrasto) risponde ad una domanda: c'è un trend quadratico tra le medie?



Quadratico $(-1 \ 1 \ 3)$

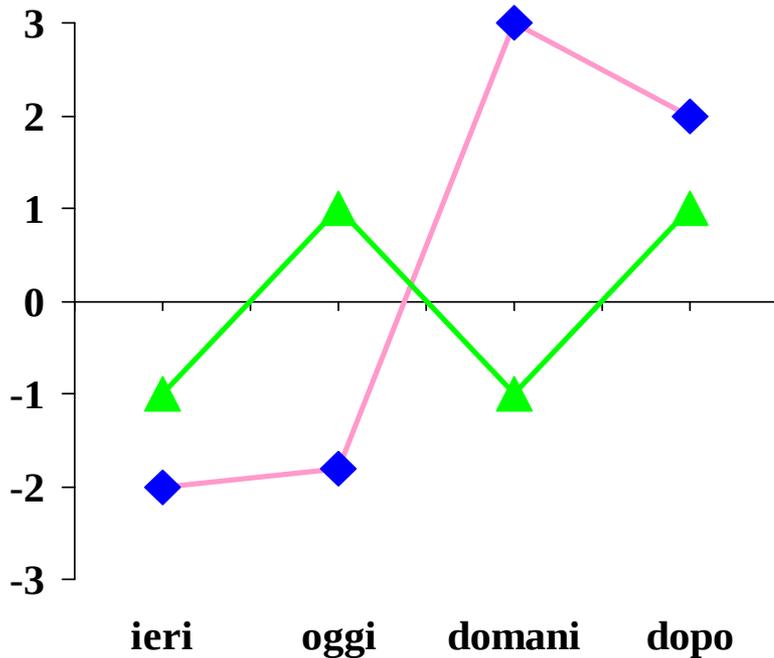
$$Q = \begin{pmatrix} -1 & 1 & 1 & -1 \end{pmatrix}$$

$$C = \begin{pmatrix} -1 & 1 & -1 & 1 \end{pmatrix}$$

C'è una tendenza nel trend ad incurvarsi?

Contrasti Polinomiali

Cioè ogni polinomio (contrasto) risponde ad una domanda: c'è un trend cubico tra le medie?



$$L = \begin{pmatrix} -3 & -1 & 1 & 3 \end{pmatrix}$$

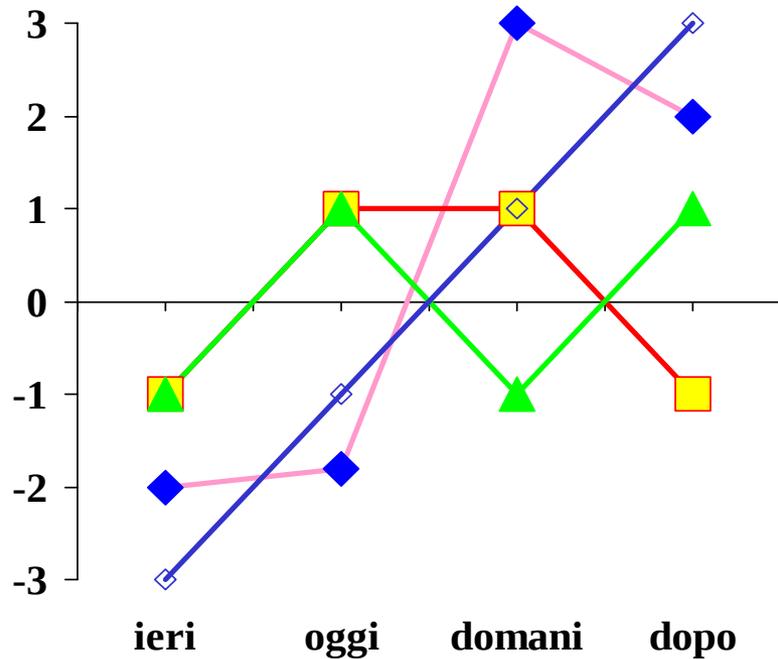
$$\text{Cubico} \begin{pmatrix} 1 & 1 & 1 & -1 \end{pmatrix}$$

$$C = \begin{pmatrix} -1 & 1 & -1 & 1 \end{pmatrix}$$

C'è una tendenza nel trend ad incurvarsi in direzioni diverse?

Contrasti Polinomiali

L'insieme dei contrasti spiega (per quanto possibile nei dati) il trend osservato nei dati



$$L = \begin{pmatrix} -3 & -1 & 1 & 3 \end{pmatrix}$$

$$Q = \begin{pmatrix} -1 & 1 & 1 & -1 \end{pmatrix}$$

$$C = \begin{pmatrix} -1 & 1 & -1 & 1 \end{pmatrix}$$

Costruire un contrasto

- In SPSS alcuni tipi di contrasti vengono stimati automaticamente dal software, altri dobbiamo specificare noi i coefficienti
- I contrasti dai noi ideati devono essere specificati mediante la sintassi di SPSS

Esperimento

Nel seguente esperimento testiamo l'effetto di un ancoraggio cognitivo sulla stima delle quantità numeriche:

- Domanda 1 a tutti i soggetti: Secondo te, le nazioni africane alle nazioni unite sono più o meno del X %.

- Domanda 2: Quante sono le nazioni africane in percentuale alle nazioni unite

- Gruppo 1: ancora 10%. Gruppo 2: ancora 80%, Nessuna ancora

Ancora

Var.
Dipendente

Tre gruppi

Condizione
sperimentale

Costruire un contrasto

Ci proponiamo ora di testare delle ipotesi specifiche sui dati “ancoraggio-numerico a tre gruppi”

Statistiche descrittive

Gruppi	N	Media	Deviazione std.
Ancora=nessuna	50	24.12	10.42
Ancora=10	50	21.14	11.22
Ancora=80	50	39.80	10.98

Costruire un contrasto

Ci proponiamo ora di testare due ipotesi specifiche sui dati “ancoraggio-numerico a tre gruppi”. Per intenderci sui pesi da dare alle medie, consideriamo ancora=10 il primo gruppo, ancora=nessuna il secondo gruppo e ancora=80 il terzo.

$$H_0: \bar{y}_{10} - \bar{y}_{nes} = 0$$

$$H_0: (\bar{y}_{10} + \bar{y}_{nes}) / 2 = \bar{y}_{80}$$

Statistiche descrittive

Gruppi	N	Media	Deviazione std.
Ancora=nessuna	50	24.12	10.42
Ancora=10	50	21.14	11.22
Ancora=80	50	39.80	10.98

Note: A green bracket labeled 'C1' spans the 'Media' column for the first two rows. A red bracket labeled 'C2' spans the 'N' column for the last two rows.

Costruire un contrasto

Ci proponiamo ora di testare due ipotesi specifiche sui dati “ancoraggio-numerico a tre gruppi”.

$$H_0: \bar{y}_{10} - \bar{y}_{nes} = 0$$

$$C_1 = (+1, -1, 0)$$

$$H_0: (\bar{y}_{10} + \bar{y}_{nes})/2 = \bar{y}_{80}$$

$$C_2 = (+1/2, +1/2, -1)$$

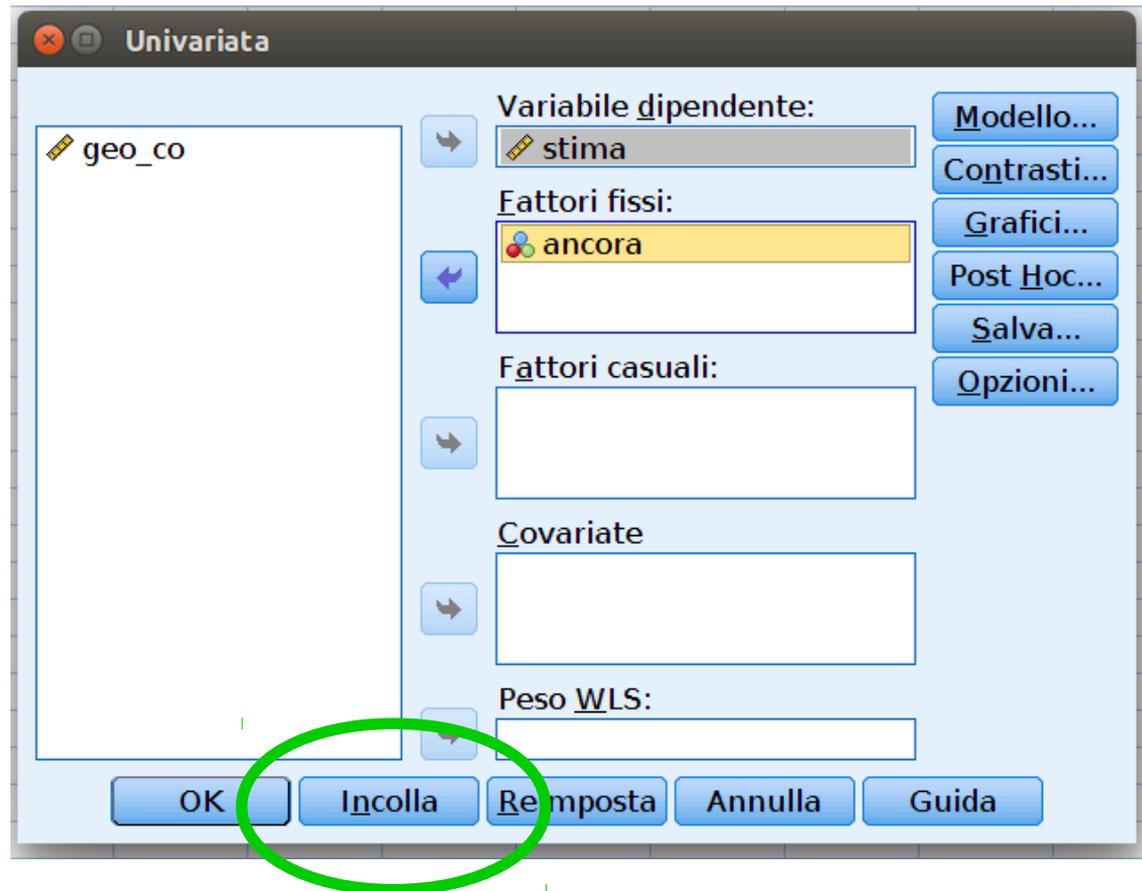
Statistiche descrittive

Gruppi	N	Media	Deviazione std.
Ancora=nessuna	50	24.12	10.42
Ancora=10	50	21.14	11.22
Ancora=80	50	39.80	10.98

Note: A green bracket labeled C1 spans the mean values of the first two rows. A red bracket labeled C2 spans the N values of the last two rows.

Costruire un contrasto

In SPSS lanciamo la ANOVA con “modello lineare generale” e poi facciamo produrre la sintassi



Costruire un contrasto

E la modifichiamo per far stimare i contrasti che ci interessano

Sintassi ottenuta

```
1  
2  
3 UNIANOVA stima BY ancora  
4 /METHOD=SSTYPE(3)  
5 /INTERCEPT=INCLUDE  
6 /CRITERIA=ALPHA(0.05)  
7 /DESIGN=ancora.  
8  
9  
10
```

Sintassi modificata

```
2  
3 UNIANOVA stima BY ancora  
4 /METHOD=SSTYPE(3)  
5 /INTERCEPT=INCLUDE  
6 /CRITERIA=ALPHA(0.05)  
7 /CONTRAST(ancora)=SPECIAL (1 -1 0)  
8 /DESIGN=ancora.  
9  
10
```

/CONTRAST(ancora)=**SPECIAL** (1 -1 0)

Var Indipendente

Coefficienti

Costruire un contrasto

E possiamo aggiungere una riga “contrast” per ogni contrasto

Sintassi ottenuta

```
1
2
3 UNIANOVA stima BY ancora
4 /METHOD=SSTYPE(3)
5 /INTERCEPT=INCLUDE
6 /CRITERIA=ALPHA(0.05)
7 /DESIGN=ancora.
8
9
10
```

Sintassi modificata

```
2
3 UNIANOVA stima BY ancora
4 /METHOD=SSTYPE(3)
5 /INTERCEPT=INCLUDE
6 /CRITERIA=ALPHA(0.05)
7 /CONTRAST(ancora)=SPECIAL (1 -1 0)
8 /CONTRAST(ancora)=SPECIAL (.5 .5 -1)
9
10 /DESIGN=ancora.
11
12
```

/CONTRAST(ancora)=SPECIAL (1 -1 0)

/CONTRAST(ancora)=SPECIAL (.5 .5 -1)

Var Indipendente

Coefficienti

Risultati

Solita tabella della ANOVA

Test di effetti tra soggetti

Variabile dipendente: stima

Origine	Somma dei quadrati di tipo III	gl	Media quadratica	F	Sign.
Modello corretto	10054.973 ^a	2	5027.487	42.441	.000
Intercetta	120586.727	1	120586.727	1017.972	.000
ancora	10054.973	2	5027.487	42.441	.000
Errore	17413.300	147	118.458		
Totale	148055.000	150			
Totale corretto	27468.273	149			

a. R-quadrato = .366 (R-quadrato adattato = .357)

Risultati

Tabelle contrasto 1

$$H_0: \bar{y}_{10} - \bar{y}_{nes} = 0 \quad C_1 = (+1, -1, 0)$$

Test sull'ipotesi personalizzati 1

Valore del
contrasto
(differenze)

Valore p. del
contrasto

Risultati del contrasto (matrice K)

		Variabile dipendente
		stima
Contrasto speciale ancora		
LI	Stima del contrasto	-3.020
	Valore ipotizzato	0
	Differenza (stima - ipotizzato)	-3.020
	Errore std.	2.177
	Sign.	.167
	95% intervallo di confidenza per differenza	
	Limite inferiore	-7.322
	Limite superiore	1.282

Risultati dei test

Variabile dipendente: stima

Origine	Somma dei quadrati	gl	Media quadratica	F	Sign.
Contrasto	228.010	1	228.010	1.925	.167
Errore	17413.300	147	118.458		

Tabella F e p.

Risultati

Tabelle contrasto 2

$$H_0 : (\bar{y}_{10} + \bar{y}_{nes}) / 2 = \bar{y}_{80} \quad C_2 = (+1/2, +1/2, -1)$$

Test sull'ipotesi personalizzati 2

Risultati del contrasto (matrice K)

		Variabile dipendente
		stima
Contrasto speciale ancora		
L1	Stima del contrasto	17.170
	Valore ipotizzato	0
	Differenza (stima - ipotizzato)	17.170
	Errore std.	1.885
	Sign.	.000
	95% intervallo di confidenza per differenza	Limite inferiore 13.445 Limite superiore 20.895

Valore del
contrasto
(differenze)

Valore p. del
contrasto

Tabella F e p.

Risultati dei test

Variabile dipendente: stima

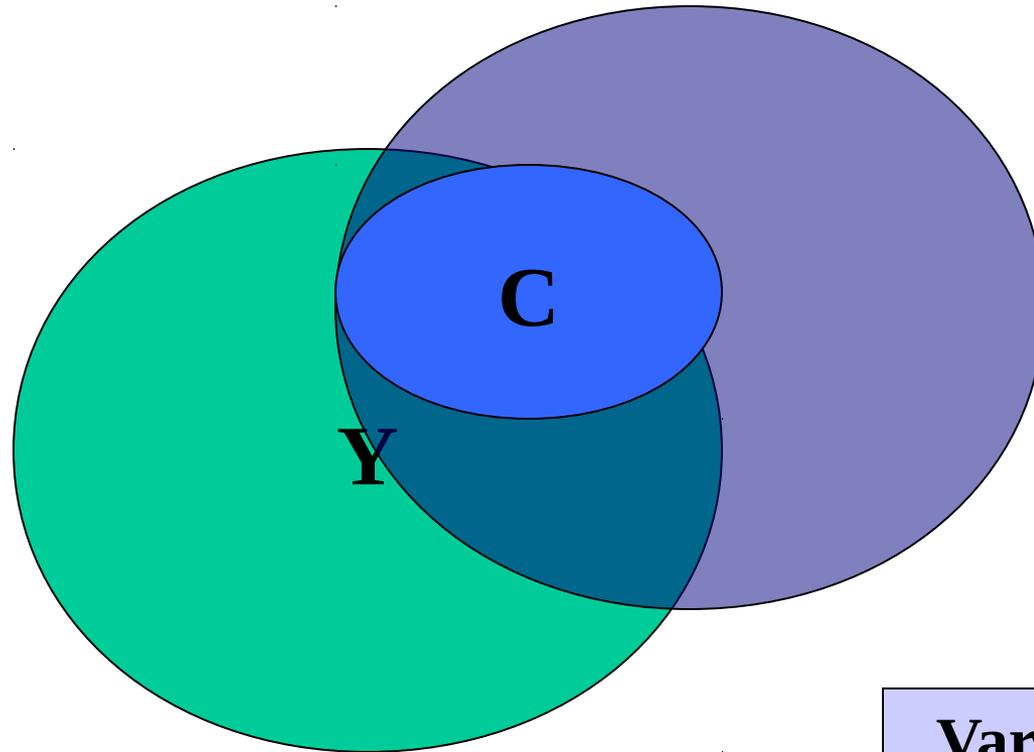
Origine	Somma dei quadrati	gl	Media quadratica	F	Sign.
Contrasto	9826.963	1	9826.963	82.957	.000
Errore	17413.300	147	118.458		

Effect sizes

**Varianza
Within**

Contrasto

**Effetto
Principale**



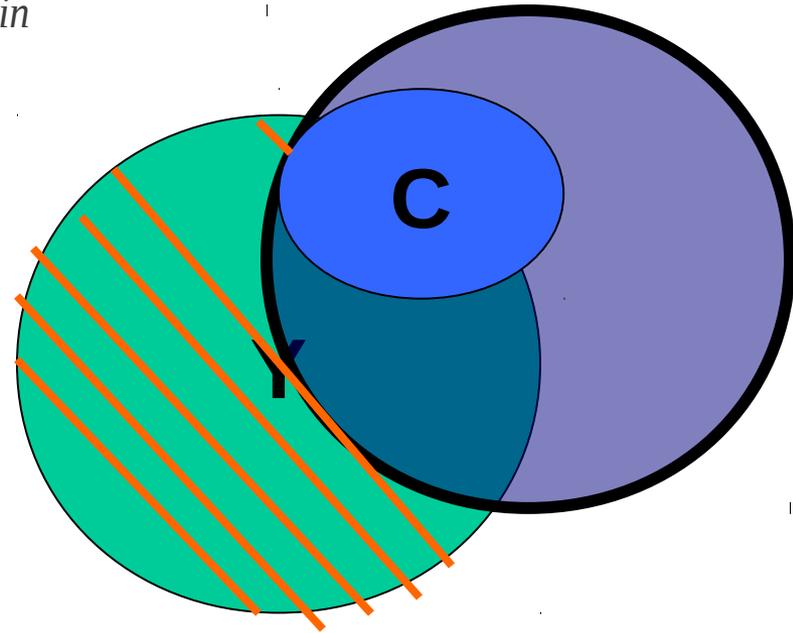
**Varianza
B**

Eta-quadro

● Prendiamo la F del contrasto $= \frac{MS_{contrast}}{MS_{within}}$

● Interpretazione:

La varianza spiegata dal contrasto dopo aver eliminato tutta la varianza spiegata dal resto del modello

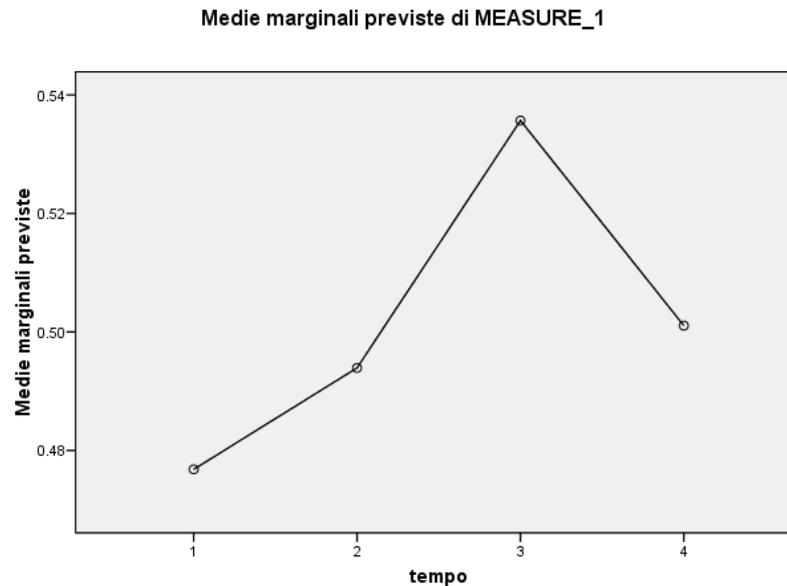


● effect size=Eta quadro

$$\frac{F_{contrast}}{F_{contrast} + df_{within}} = pr^2 = p\eta^2$$

Esempio dati

Disegno di ricerca: Abbiamo un esperimento in cui un tre set di parole vengono somministrate ad un campione in sequenza (1 2 3 e 4) sotto nel tempo. Si vuole stabilire se vi sia un peggioramento della performance mnemonica (misurata in proporzione di errori)



Esempio

Risultati: Abbiamo un effetto significativo del fattore tempo, ora vogliamo stabilire la forma dell'andamento delle medie nel tempo

Test degli effetti entro soggetti

Misura: MEASURE_1

Sorgente		Somma dei quadrati Tipo III	df	Media dei quadrati	F	Sig.
tempo	Assumendo la sfericità	.367	3	.122	8.679	.000
	Greenhouse-Geisser	.367	2.335	.157	8.679	.000
	Huynh-Feldt	.367	2.364	.155	8.679	.000
	Limite inferiore	.367	1.000	.367	8.679	.004
Errore(tempo)	Assumendo la sfericità	8.408	597	.014		
	Greenhouse-Geisser	8.408	464.674	.018		
	Huynh-Feldt	8.408	470.530	.018		
	Limite inferiore	8.408	199.000	.042		

Esempio

SPSS default: SPSS fornisce automaticamente i contrasti polinomiali

Test dei contrasti entro soggetti

Misura: MEASURE_1

Sorgente	tempo	Somma dei quadrati Tipo III	df	Media dei quadrati	F	Sig.
tempo	Lineare	.131	1	.131	5.415	.021
	Quadratico	.134	1	.134	12.505	.001
	Cubico	.102	1	.102	13.858	.000
Errore(tempo)	Lineare	4.816	199	.024		
	Quadratico	2.129	199	.011		
	Cubico	1.463	199	.007		

Come li interpretiamo

Esempio

SPSS default: SPSS fornisce automaticamente i contrasti polinomiali

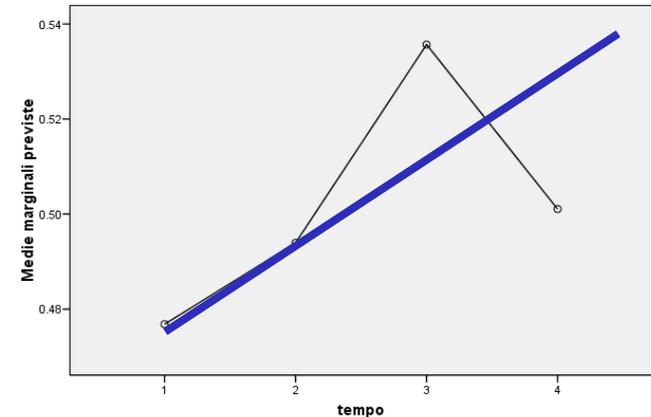
Test dei contrasti entro soggetti

Misura: MEASURE_1

Sorgente	tempo	Somma dei quadrati Tipo III	df	Media dei quadrati	F	Sig.
tempo	Lineare	.131	1	.131	5.415	.021
	Quadratico	.134	1	.134	12.505	.001
	Cubico	.102	1	.102	13.858	.000
Errore(tempo)	Lineare	4.816	199	.024		
	Quadratico	2.129	199	.011		
	Cubico	1.463	199	.007		

Se il trend lineare è significativo, le medie tendono ad aumentare (diminuire nel tempo)

Medie marginali previste di MEASURE_1



Esempio

SPSS default: SPSS fornisce automaticamente i contrasti polinomiali

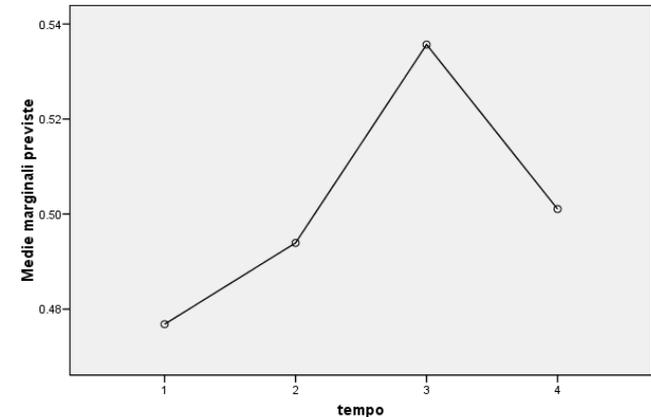
Test dei contrasti entro soggetti

Misura: MEASURE_1

Sorgente	tempo	Somma dei quadrati Tipo III	df	Media dei quadrati	F	Sig.
tempo	Lineare	.131	1	.131	5.415	.021
	Quadratico	.134	1	.134	12.505	.001
	Cubico	.102	1	.102	13.858	.000
Errore(tempo)	Lineare	4.816	199	.024		
	Quadratico	2.129	199	.011		
	Cubico	1.463	199	.007		

Se il trend quadratico è significativo, le medie tendono ad aumentare prima e diminuire dopo

Medie marginali previste di MEASURE_1



Esempio

SPSS default: SPSS fornisce automaticamente i contrasti polinomiali

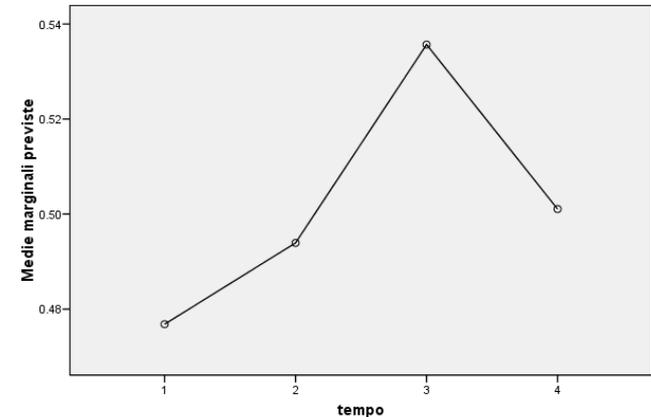
Test dei contrasti entro soggetti

Misura: MEASURE_1

Sorgente	tempo	Somma dei quadrati Tipo III	df	Media dei quadrati	F	Sig.
tempo	Lineare	.131	1	.131	5.415	.021
	Quadratico	.131	1	.131	12.505	.001
	Cubico	.102	1	.102	13.858	.000
Errore(tempo)	Lineare	4.816	199	.024		
	Quadratico	2.129	199	.011		
	Cubico	1.463	199	.007		

Se il trend cubico è significativo, il grafico delle medie contiene una parte a forma di S

Medie marginali previste di MEASURE_1



Interpretazione alternativa

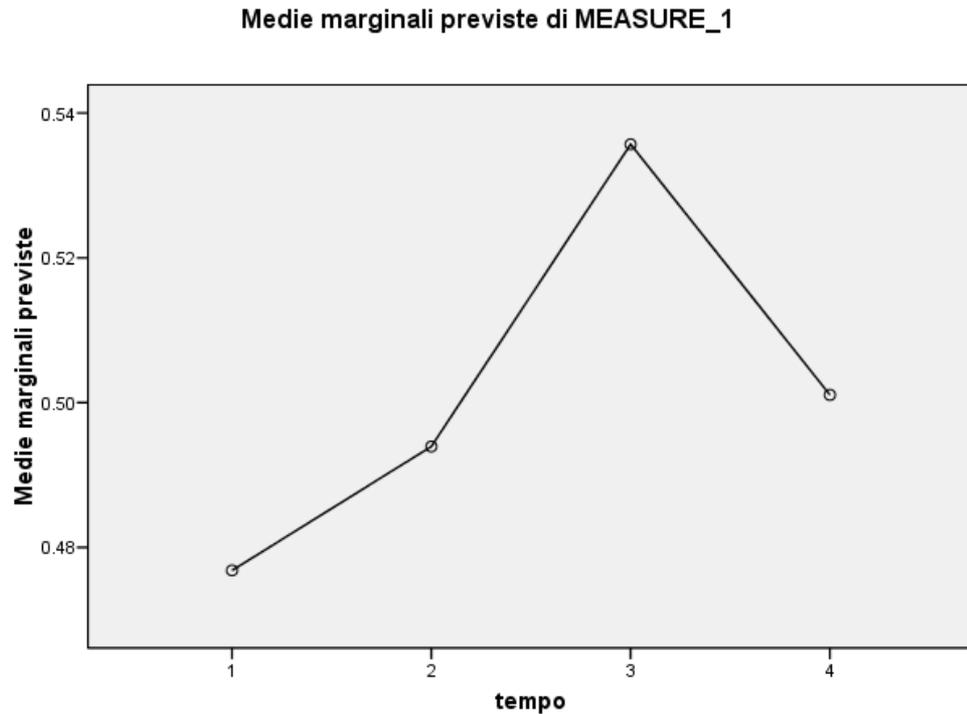
Ricordando i persi:

I contrasti si possono anche interpretare non in base al tempo, ma in base alle misure confrontate

$$L = \begin{pmatrix} -3 & -1 & 1 & 3 \end{pmatrix}$$

$$Q = \begin{pmatrix} -1 & 1 & 1 & -1 \end{pmatrix}$$

$$C = \begin{pmatrix} -1 & 1 & -1 & 1 \end{pmatrix}$$



Interpretazione alternativa

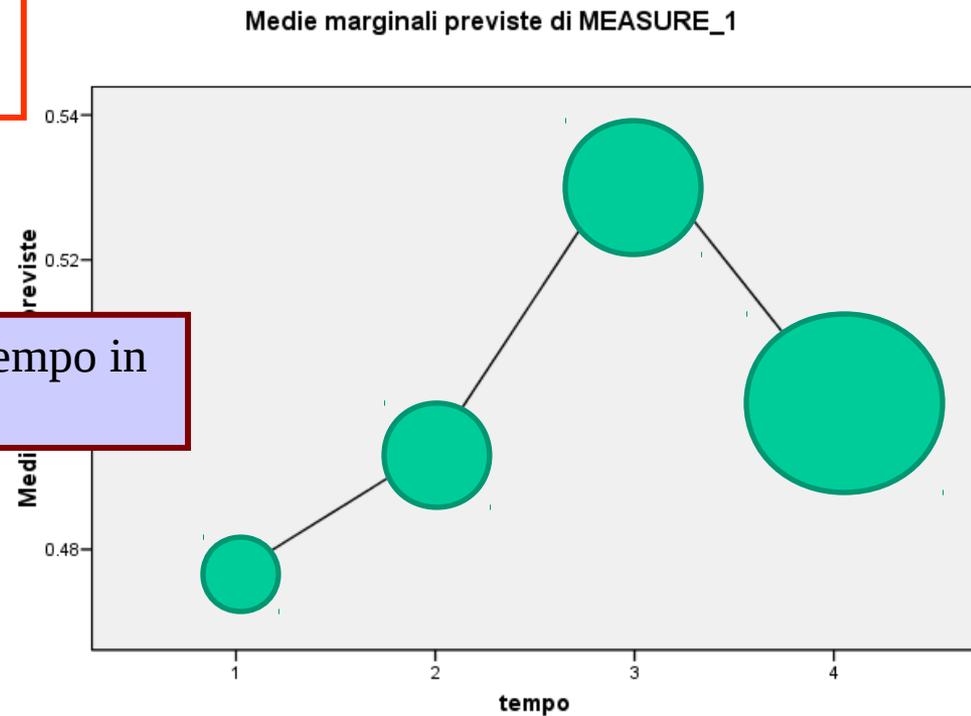
Ricordando i persi:

Lineare

$$L = \begin{pmatrix} -3 & -1 & 1 & 3 \end{pmatrix}$$

I contrasti si possono anche interpretare non in base al tempo, ma in base alle misure confrontate

In media, abbiamo che per ogni tempo in più le medie aumentano?



Interpretazione alternativa

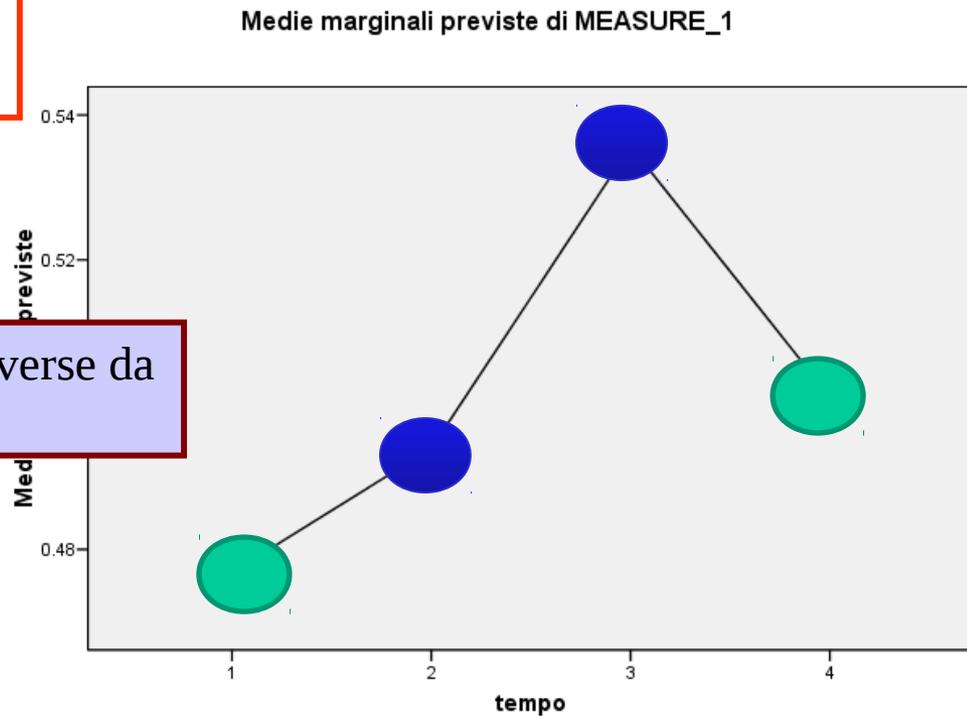
Ricordando i persi:

Quadratico

$$Q = \begin{pmatrix} -1 & 1 & 1 & -1 \end{pmatrix}$$

I contrasti si possono anche interpretare non in base al tempo, ma in base alle misure confrontate

In media, le medie T1+T4 sono diverse da T2+T3



Interpretazione alternativa

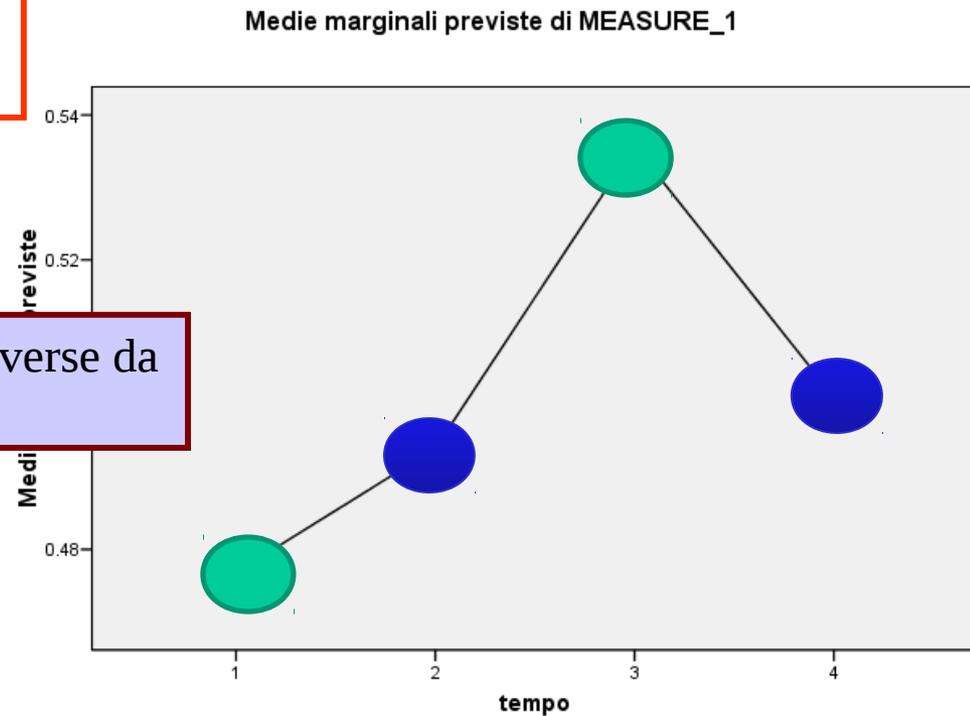
Ricordando i persi:

Cubico

$$C = \begin{pmatrix} -1 & 1 & -1 & 1 \end{pmatrix}$$

I contrasti si possono anche interpretare non in base al tempo, ma in base alle misure confrontate

In media, le medie T1+T3 sono diverse da T2+T4



Test post hoc

Nessuna ipotesi

Ci proponiamo ora di testare tutte le possibili differenze tra gruppi sui dati “ancoraggio- numerico a tre gruppi”.

Statistiche descrittive

Gruppi	N	Media	Deviazione std.
Ancora=nessuna	50	24.12	10.42
Ancora=10	50	21.14	11.22
Ancora=80	50	39.80	10.98

Confronti Post-Hoc

- I confronti post-hoc (a posteriori) servono a trovare le differenze tra i gruppi, presi a due a due.

Ipotesi nulla ANOVA

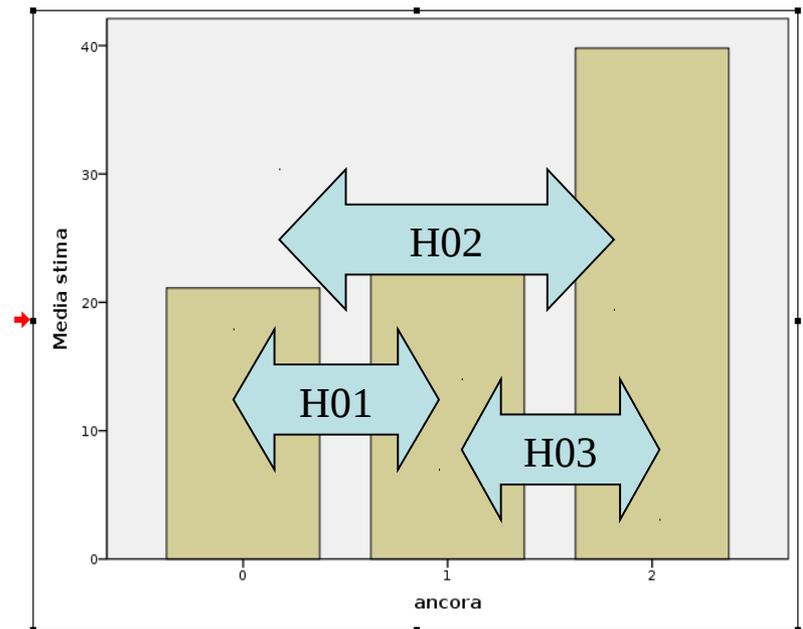
$$H_0 : \mu_1 = \mu_2 = \mu_3$$

Ipotesi nulle Post-Hoc

$$H_{01} : \mu_1 = \mu_2$$

$$H_{02} : \mu_1 = \mu_3$$

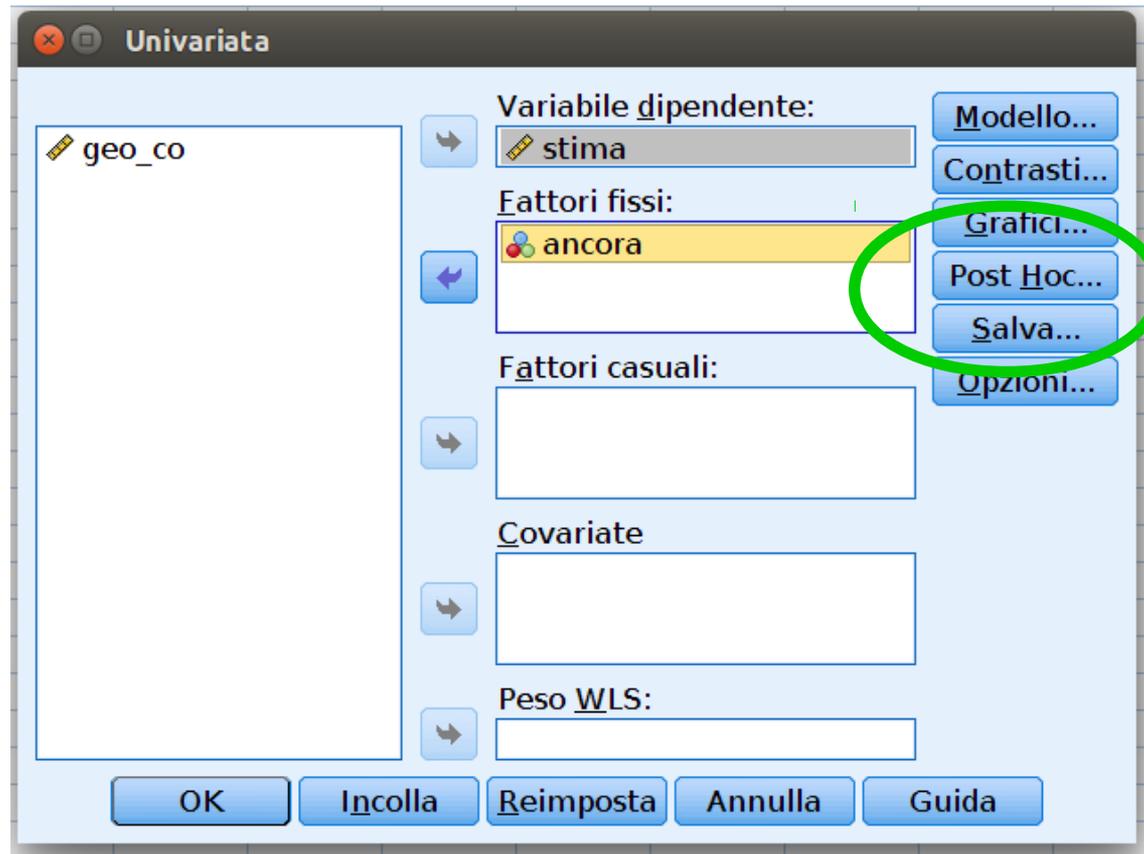
$$H_{03} : \mu_2 = \mu_3$$



Scriviamo le medie come μ (mu) in quanto l'ipotesi riguarda la popolazione

Post-hoc

In SPSS lanciamo la ANOVA con “modello lineare generale” e selezioniamo Post Hoc



Confronti Post-Hoc: SPSS

- Esistono decine di confronti post-hoc!!!

Univariata: Confronti multipli post hoc per medie osservate

Fattori
ancora

Test post hoc per:
ancora

Varianze uguali presunte

LSD S-N-K Waller-Duncan
 Bonferroni Tukey Rapporto dell'errore tipo I/tipo II: 100
 Sidak Tukey's-b Dunnett
 Scheffe Duncan Categoria di controllo: Ultima
 R-E-G-W-F GT2 di Hochberg Test
 R-E-G-W-Q Gabriel bilaterale < Controllo > Controllo

Varianze uguali non presunte

T2 di Tamhane T3 di Dunnett Games-Howell C di Dunnett

Continua Annulla Guida

Finestra Post-Hoc

Tutti, in una maniera o in un'altra, cercano di testare le differenze tra le medie a due a due
Noi usiamo "Tukey" e "REGWG"

Post-Hoc: Output

- Tabella post hoc

Confronti multipli

Variabile dipendente: stima

	(I) ancora	(J) ancora	Differenza della media (I-J)	Errore std.	Sign.	Intervallo di confidenza 95%	
						Limite inferiore	Limite superiore
HSD di Tukey	0	1	-3.02	2.177	.350	-8.17	2.13
		2	-18.68*	2.177	.000	-23.83	-13.53
	1	0	3.02	2.177	.350	-2.13	8.17
		2	-15.66*	2.177	.000	-20.81	-10.51
	2	0	18.68*	2.177	.000	13.53	23.83
		1	15.66*	2.177	.000	10.51	20.81

Si basa sulle medie osservate.

Interpretiamo la significatività dei confronti

Confronti Post-Hoc: Confronti multipli

- Nei confronti multipli ogni gruppo è testato per le differenze con ogni altro

Confronti multipli

Significatività

No Ancora non diverso da Ancora 10

Confronti multipli

		stima		Differenza della media (I-J)	Errore std.	Sign.	Intervallo di confidenza 95%	
(I) ancora	(J) ancora						Limite inferiore	Limite superiore
HSD di Tukey	0	1		-3.02	2.177	.350	-8.17	2.13
		2		-18.68*	2.177	.000	-23.83	-13.53
	1	0		3.02	2.177	.350	-2.13	8.17
		2		-15.66*	2.177	.000	-20.81	-10.51
	2	0		18.68*	2.177	.000	13.53	23.83
		1		15.66*	2.177	.000	10.51	20.81

Si basa sulle medie osservate.

Significatività

Ancora 80 è diversa da tutti gli altri gruppi?

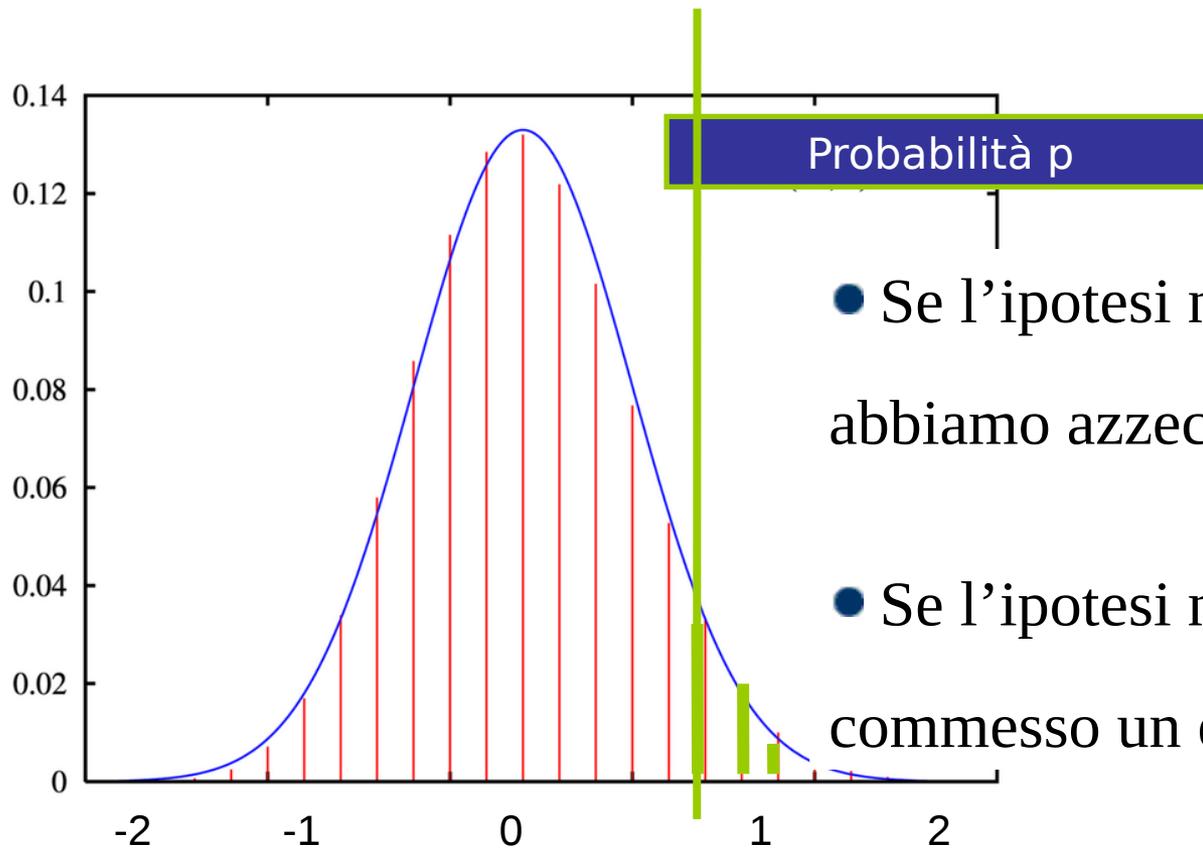
Confronti Post-Hoc: Domanda

- In generale, non siamo interessati ai calcoli specifici di ogni confronto post-hoc.
- In pratica ci basta capire come interpretarli
- Vogliamo ora capire perchè non possiamo semplicemente fare tutti i confronti usando dei semplici t-test

- Ricordiamo che quando facciamo un test inferenziale stimiamo la probabilità di commettere un errore rifiutando l'ipotesi nulla quando è vera

VALORE-P

- Il valore p indica il rischio che noi prendiamo quando affermiamo che l'ipotesi nulla è falsa



- Se l'ipotesi nulla è falsa, ci abbiamo azzeccato
- Se l'ipotesi nulla è vera, abbiamo commesso un errore, detto del **Tipo I**

Il problema del caso

- Se l'ipotesi nulla è vera (non ci sono realmente delle differenze) e rifiutiamo l'ipotesi nulla ogni volta che $p < .05$ (*alpha critico*), alla lunga commettiamo il 5% di errore
- Il 5% va bene, di più no!
- Ciò è equivalente alla situazione in cui compriamo un biglietto di una lotteria in cui ci sono 20 numeri in totale. Un numero su 20 esce, ogni biglietto ha un numero. Abbiamo il 5% di probabilità di essere estratti e 95% di non essere estratti
- Se il biglietto è estratto, le nostre conclusioni sono sbagliate!

Estrazioni multiple

- Cosa succede se partecipiamo a due estrazioni
- La probabilità di essere estratto aumenta! Più estrazioni si fanno, più aumenta la probabilità di essere estratti
- Cioè, anche se il biglietto ha il 5% di chance di essere estratto in una estrazione, facendo varie estrazioni la probabilità di essere estratto aumenterà
- In maniera equivalente, più test facciamo, più aumenta la probabilità di rifiutare l'ipotesi nulla anche se essa è vera
- Cioè aumenta la probabilità di sbagliare.

Test multipli

- A seconda di quanti test facciamo, la probabilità di ottenere almeno una differenza significativa per caso

Criterio usato

p. Nessun test significativo

p. Almeno uno significativo

Un test solo

$$\alpha = .05$$

$$ns = .95$$

$$\alpha_{vero} = .05$$

Due test

$$ns = .95 \cdot .95 = .9025$$

$$\alpha_{vero} = 1 - .9025 = .09275$$

Tre test

$$ns = .95^3 = .8573$$

$$\alpha_{vero} = 1 - .8573 = .1423$$

Test multipli

- In generale, facendo K test, se rifiutiamo l'ipotesi nulla per $p < .05$, il nostro errore diventa sempre maggiore e maggiore di 0.05

Criterio usato

$$\alpha = .05$$

$$\alpha = C$$

p. Nessun test significativo

$$ns_k = .95^k$$

$$ns_k = C^k$$

p. di errore

$$\alpha_{\text{vero}} = 1 - .95^k$$

$$\alpha_{\text{vero}} = 1 - C^k$$

Confronti post-hoc

- In generale, i vari confronti post-hoc cercano di calcolare le probabilità associate ai vari confronti in modo tale che

p. di errore

$$\alpha_{\text{vero}} = 1 - C^k \Rightarrow 0.05 = \alpha$$

Cioè cercando di fissare la probabilità di ottenere almeno un test significativo quando la differenza sono 0, uguale a quella di come se facessimo un test solo

In pratica, i vari test post-hoc usano vari espedienti per controllare questa probabilità ai valori corretti

Il test post-hoc più conservativo

DISUGUAGLIANZA DI BONFERRONI:

Dati c confronti *post hoc*,

probabilità che almeno uno sia significativo per caso $\leq c * \alpha_c$

dove α_c è il valore che adottato per decidere se il singolo confronto è significativo.



Scelgo il valore $\alpha_c = \alpha / c$

Esempio: se il nr di confronti totale è 6 e voglio che il valore complessivo $\alpha = .05$,

per ciascun confronto giudico la differenza come significativa solo se $p < (.05 / 6)$, ossia se $p < .0083$.

= criterio di Bonferroni

Estrazioni multiple

- Quando facciamo dei confronti multipli non pianificati dobbiamo usare una correzione
- Usiamo Tuckey or REGWQ sui nostri dati
- Se abbiamo dei test ma non abbiamo i dati (analisi pubblicate da altri) possiamo sempre usare il criterio Bonferroni in quanto si applica sul criterio di significatività

Fine

Fine della Lezione IX

