Il modello lineare generale, misto e generalizzato

Introduzione

Marcello Gallucci
Università Milano-Bicocca

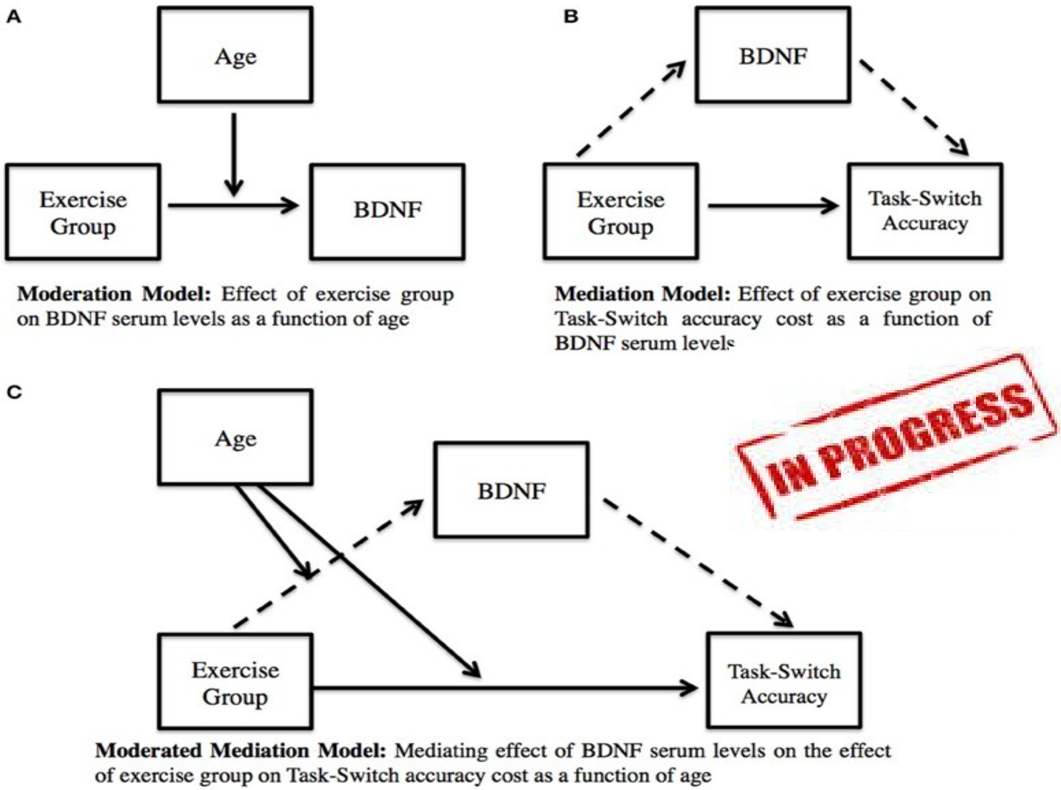
Preludio

Preludio

- Il corso ha come scopo di introdurre ed approfondire tre dei più importanti modelli statistici utilizzati in psicologia
  - Il modello lineare generale (*regressione/anova*)
  - Il modello lineare misto (*random coefficients models, multilevel models*)
  - Il modello lineare generalizzato (l*ogistica, poisson, multinomiale*)
- Rivedremo insieme anche le caratteristiche di base di ognuno di questi modelli generali

Preludio

- Con particolare enfasi all'applicazione di tali modelli statistici nello studio della mediazione e della moderazione.



Software

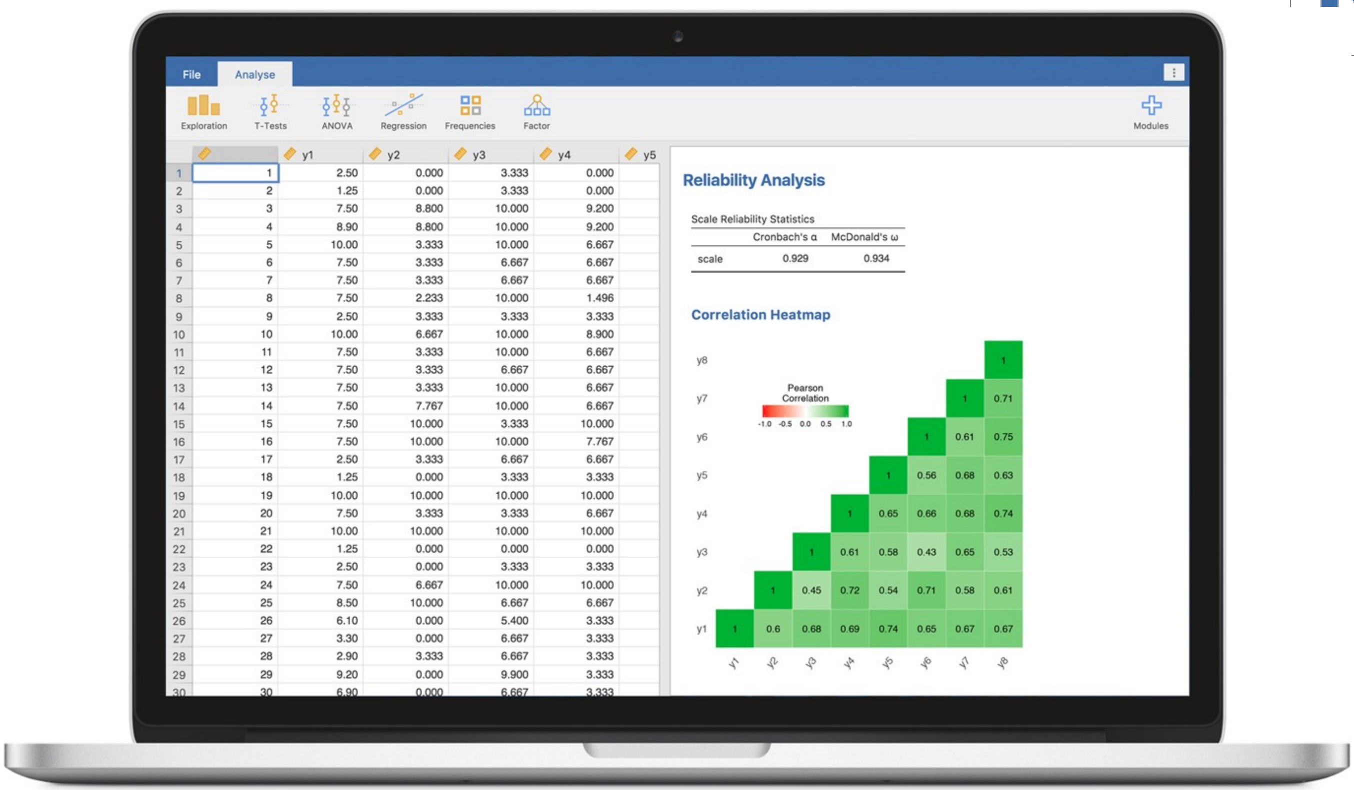
R

jamovi Stats. Open. Now.

SPSS

jamovi

www.jamovi.org



jamovi Stats. Open. Now.

Dati e file

https://github.com/mcfanda/binaries

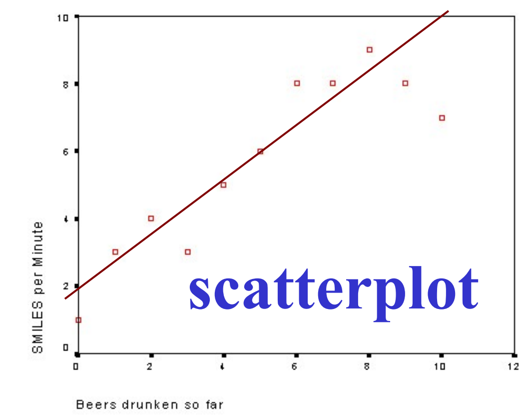
https://github.com/mcfanda/binaries/Slides

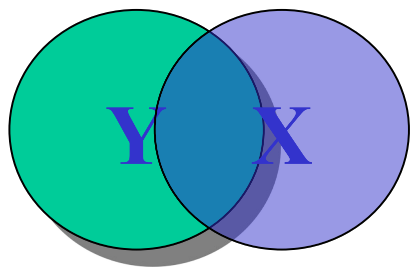
https://github.com/mcfanda/binaries/data

Introduzione

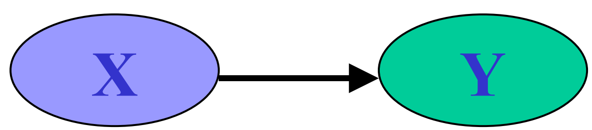
Introduzione

Un semplice **modello statistico** è una rappresentazione efficiente e compatta dei dati raccolti per descrivere un fenomeno empirico

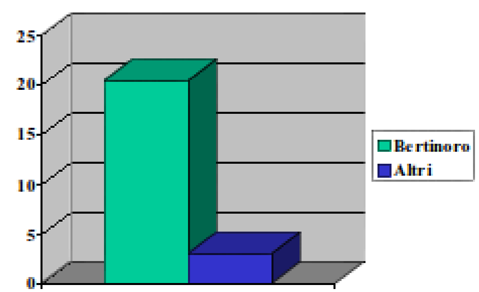




Path Diagram

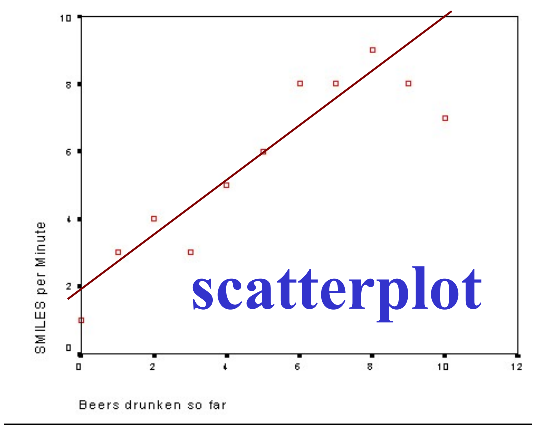


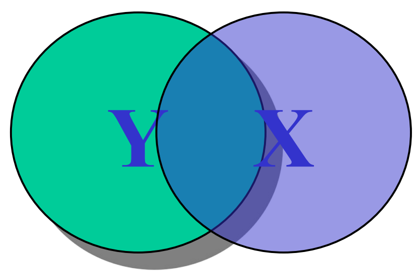
Differenze medie



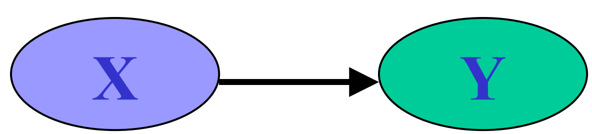
Introduzione

La maggior parte delle tecniche statistiche che conosciamo (e incontreremo) definiscono un **modello statistico** delle **relazioni** fra variabili di interesse

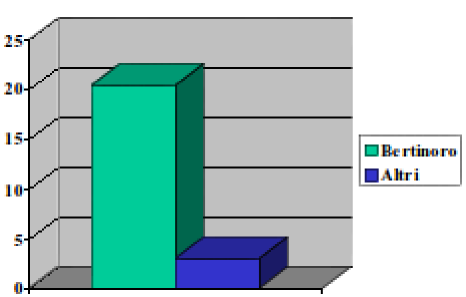




Path Diagram



Differenze medie

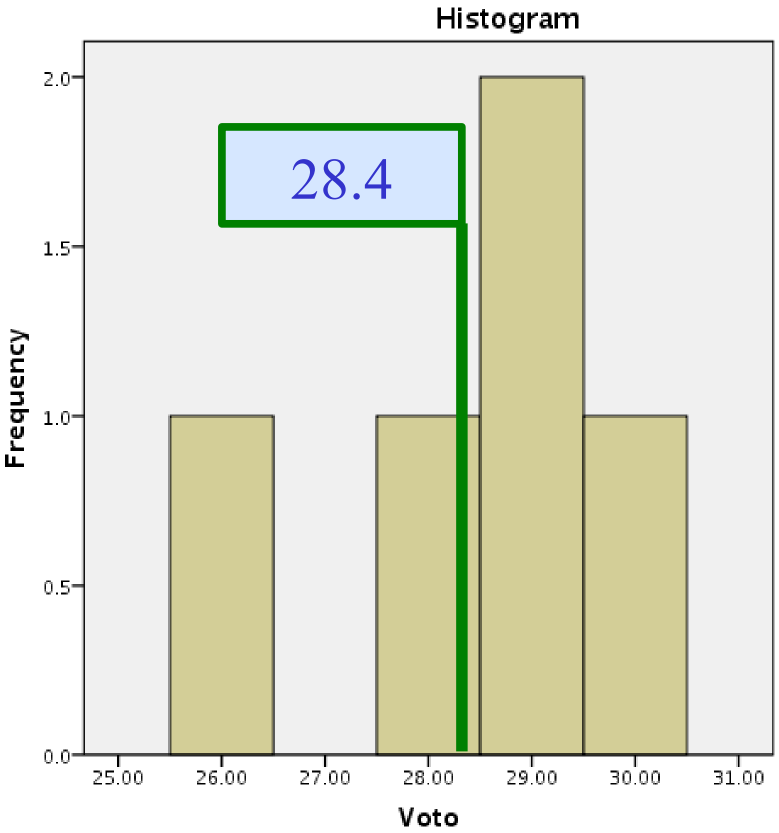


Esempio: la media

Q: “Come vanno gli studenti al mio corso?”

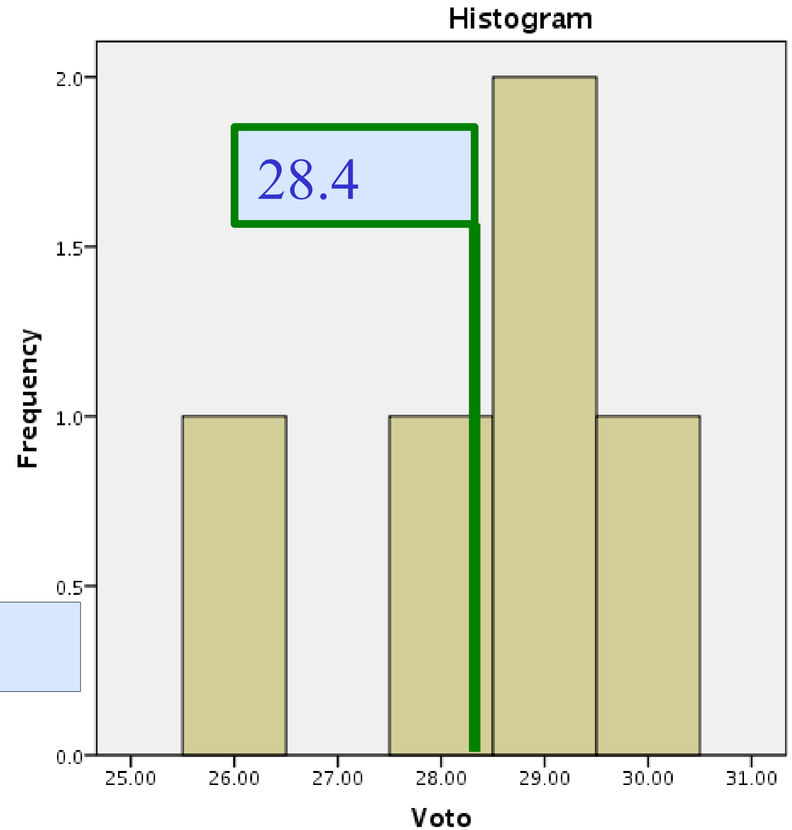
R: “Hanno una media del 28.4”

$$\frac{\sum_i X_i}{N} = \bar{X}$$



Introduzione

- Il modello statistico e la rappresentazione che ne facciamo serve (tra l'altro) a tre scopi:
  - Descrizione efficiente e compatta
  - Predizione del futuro
  - Inferenza sulla popolazione

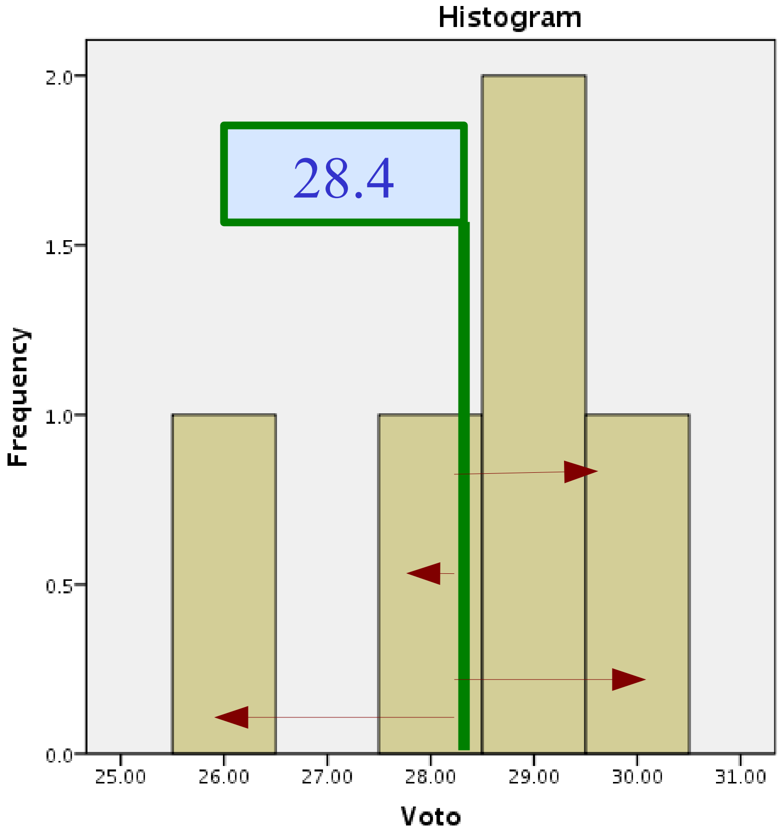


Cioè: comprensione del fenomeno

Errore di approssimazione

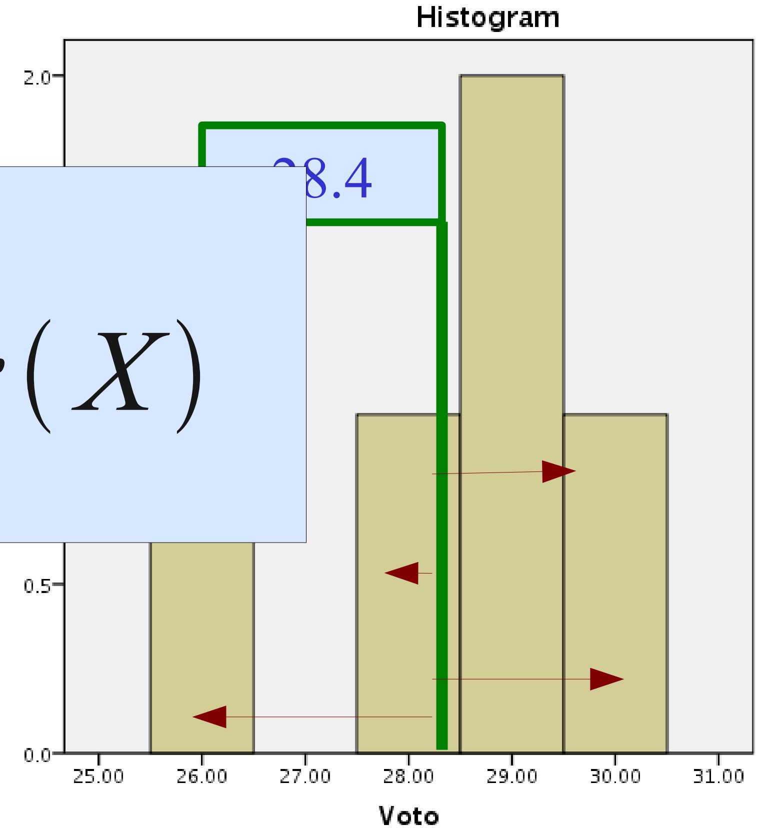
Come tutte le rappresentazioni compatte ed efficienti, anche quella statistica è una approssimazione dei dati rappresentati

Se, per semplificare, diremo che la performance è di 28.4, mis-rappresenteremo alcuni dei voti effettivi



Errore di approssimazione

Calcolando questo errore per ogni caso (ogni studente), elevandolo al quadrato (sbagliare in più o in meno è uguale) e facendo la media per ogni caso, quantifichiamo l'errore medie associato alla media

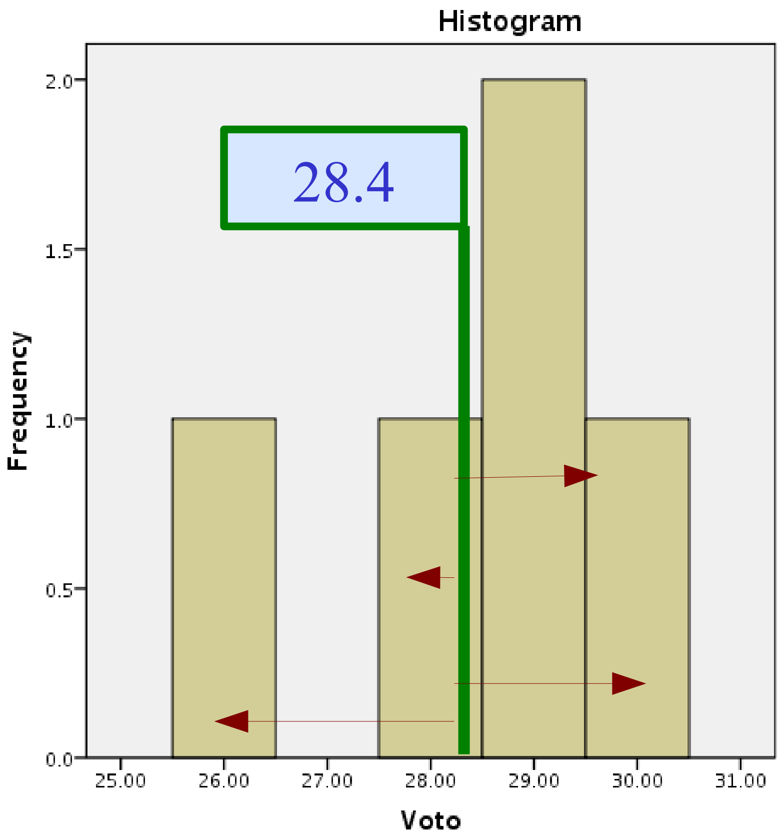


$$\frac{\sum_i (X_i - \bar{X})^2}{N-1} = Var(X)$$

Inferenza statistica

Il modello statistico è associato ad una serie di test inferenziali che ci consentono di trarre conclusioni sulla popolazione di riferimento

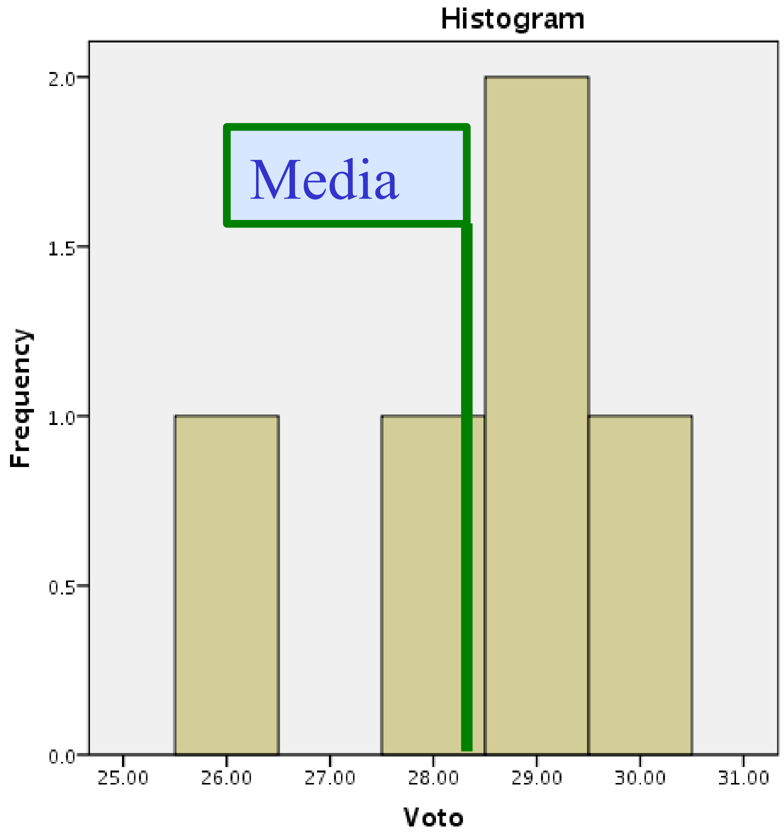
$$\frac{\bar{X}-\mu}{\sqrt{\frac{Var(X)}{N}}}=ttest$$



Modello statistico

Il modello statistico sarà una buona rappresentazione dei dati se:

- I parametri sono modellati correttamente
- Gli errori sono modellati correttamente
- La struttura dei dati è rispettata



Scegliere un modello statistico

Per costruire un corretto modello statistico dei nostri dati dobbiamo sapere una serie di cose:

- Cosa ci serve il modello (lo scopo dell'analisi)
- Che tipo di variabili abbiamo
- Che tipo di relazioni vogliamo studiare
- Quali sono le unità di misurazioni dei dati
- Come sono strutturati i nostri dati

Scegliere un modello statistico

Per costruire un corretto modello statistico dei nostri dati dobbiamo sapere una serie di cose:

- Cosa ci serve il modello (lo scopo dell'analisi)
- Che tipo di variabili abbiamo
- **Che tipo di relazioni vogliamo studiare**
- Quali sono le unità di misurazioni dei dati
- Come sono strutturati i nostri dati

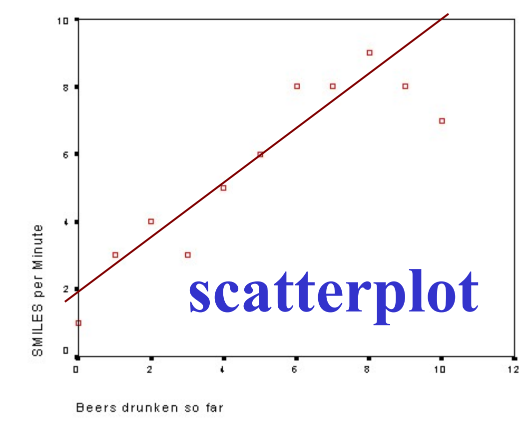
Struttura delle relazioni

- La relazione più semplice che conosciamo è la correlazione fra due variabili

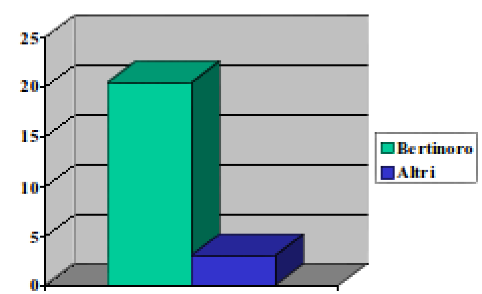
Modello 1

$$X \xrightarrow{r} Y$$

- In cui la X può essere o continua o categorica



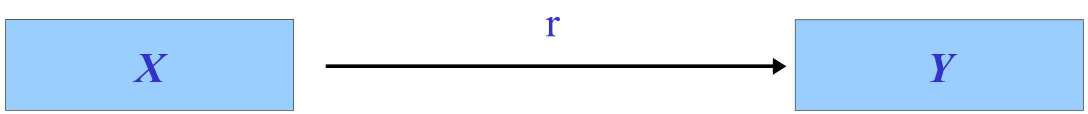
Differenze medie



Struttura delle relazioni

- Una relazione semplice tra X e Y indica che

Modello 1



- Le due variabili si muovono insieme: al cambiare dei valori di X cambiano (in media) i valori di Y

- X è un predittore di Y: sapendo i valori di X possiamo stimare i valori di Y

- X ha un effetto su Y: modificando i valori di X possiamo modificare i valori di Y (*)

Struttura causale delle relazioni

- X ha un effetto su Y: modificando i valori di X possiamo modificare i valori di Y (*)

Modello 1

$$X \xrightarrow{r} Y$$

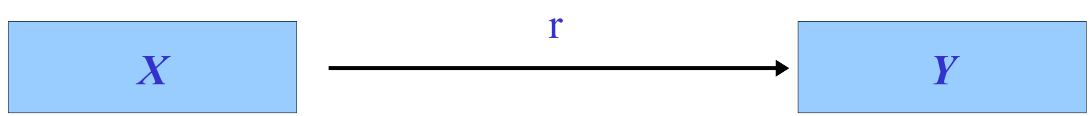
Modello 1b

$$Y \xrightarrow{r} X$$

- Una relazione statistica **non prova mai una relazione causale**: l'ipotesi causale va giustificata con:
    - Metodo sperimentale
    - Metodi temporali (longitudinali)
    - Teoria

Struttura delle relazioni

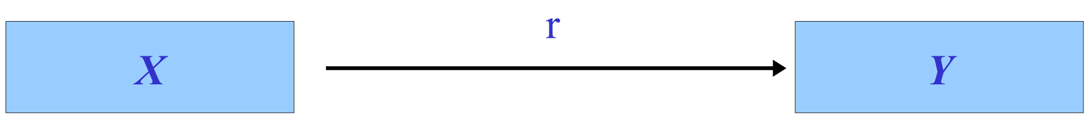
- Le relazioni possibili diventano più interessanti strutturalmente quando siamo in presenza di tre o più variabili



- Una terza variabile può intervenire in vari modi nella relazione tra una variabile indipendente (IV) ed una dipendente (DV)

Mediazione e Moderazione

- L'analisi della mediazione e della moderazione servono a comprendere come una (o più) terze variabili intervengono nella relazione tra due (o più) variabili.



- Attengono cioè allo studio della struttura delle relazioni: come le relazioni tra X e Y sono influenzate da Z

Esempio

- In un esperimento i partecipanti, divisi in due gruppi sperimentali, sono sottoposti a prime di "might" vs "morality" (*prime*). Poi svolgono un compito cooperativo in cui possono cooperate con diversa intensità (BEH). Essendo la cooperazione associata sia a valori individuali che alle aspettative sull'opponente, le aspettative di cooperazione dell'altro sono state chieste ad ogni soggetto (EXP), ed una misura continua di Social Value Orientation (SVO) è stata presa, con valori alti corrispondenti a maggiore tratto di cooperatività

- L'ipotesi iniziale è che il prime "morality" induca maggiore cooperatività

$$\boxed{\textit{PRIME}} \xrightarrow{c} \boxed{\textit{BEH}}$$

Dataset "medmod",
Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is meditated and when mediation is moderated. JPSP, 89(6), 852-863

Capire meglio gli effetti

- Supponiamo di aver trovato una relazione tra PRIME e BEH.

PRIME $\xrightarrow{c}$ ***BEH***

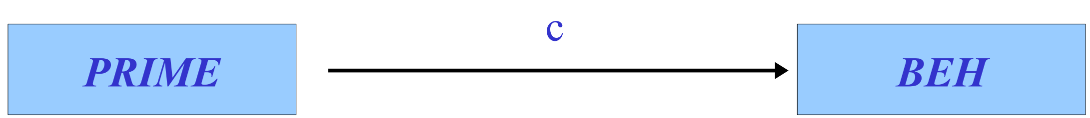
- L'analisi (logica per ora) della mediazione e della moderazione ci aiutano a capire meglio questa relazione grazie all'intervento di altre variabili, cioè EXP (aspettative di cooperazione) e SVO (il tratto di cooperatività)

EXP

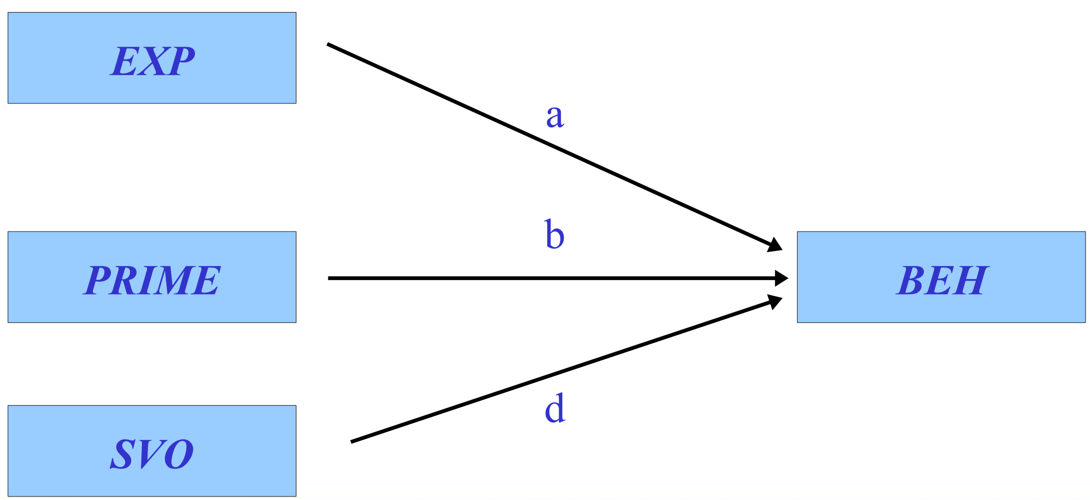
SVO

Esempio

- Una possibilità è studiare la relazione tra le nostre variabili nel **contesto** di altre variabili predittrici

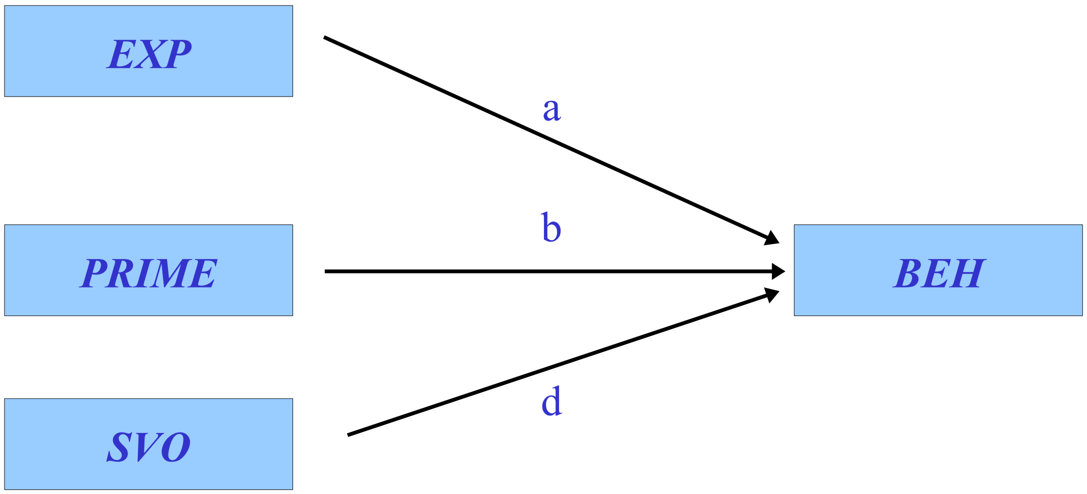


- E domandarci se e come le variabili indipendenti predicono/spiegano la dipendente al netto degli effetti delle altre



Modelli multipli

- In un **modello lineare multiplo** esaminiamo e stimiamo gli effetti di due o più variabili indipendenti, ogni effetto **al netto degli effetti** delle altre variabili



- Ogni variabile indipendente è una **covariata** allo stesso livello delle altre variabili indipendenti

Esempio

- Supponiamo di aver trovato una relazione tra PRIME e BEH (comportamento cooperativo).

$$\textit{PRIME} \xrightarrow{c} \textit{BEH}$$

- Possiamo domandarci perché PRIME abbia un effetto su BEH
  - Possiamo ipotizzare che coloro che sono *primed* con morality sviluppino delle aspettative più alte di cooperazione dell'altro rispetto a chi è *primed* con “might”

$$\textit{PRIME} \xrightarrow{a} \textit{EXP}$$

Quesito sul perchè

- Possiamo domandarci perché PRIME abbia un effetto su BEH

  • Possiamo ipotizzare che coloro che sono *primed* con morality sviluppino delle aspettative più alte di cooperazione dell'altro rispetto a chi è *primed* con "might"

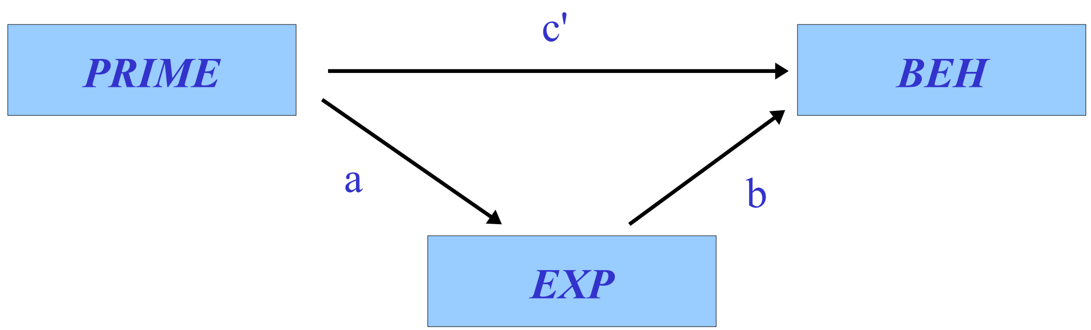
$$\textbf{\textit{PRIME}} \xrightarrow{\quad\quad a \quad\quad} \textbf{\textit{EXP}}$$

  • E che avere delle aspettative di maggior cooperazione dell'altro porti a maggiore cooperazione nel partecipante

$$\textbf{\textit{EXP}} \xrightarrow{\quad\quad\quad\quad} \textbf{\textit{BEH}}$$

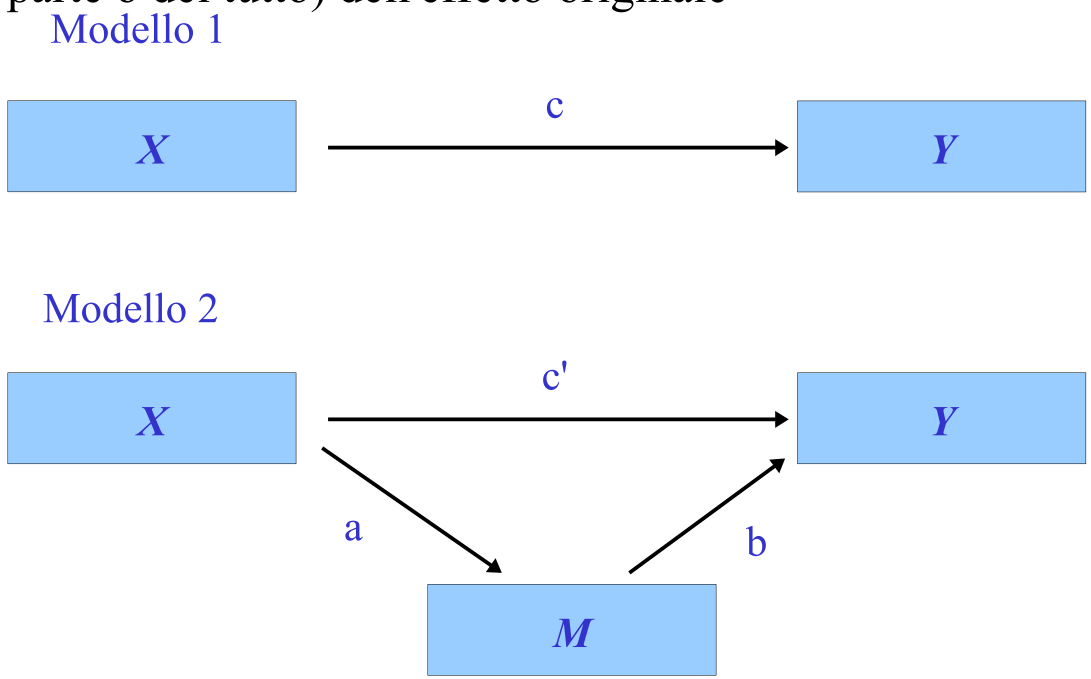
Esempio

- E dunque, uno dei motivi per cui PRIME ha un effetto sul comportamento (BEH), è che PRIME aumenta le aspettative (EXP), e le aspettative aumentano la cooperazione (BEH)



In generale

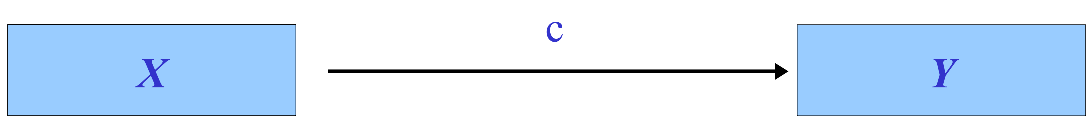
- In presenza di una relazione tre una IV (X) e una VD (Y), possiamo domandarci se uno dei motivi per cui osserviamo un effetto è l'intervento di una terza variabile M, che è responsabile (in parte o del tutto) dell'effetto originale



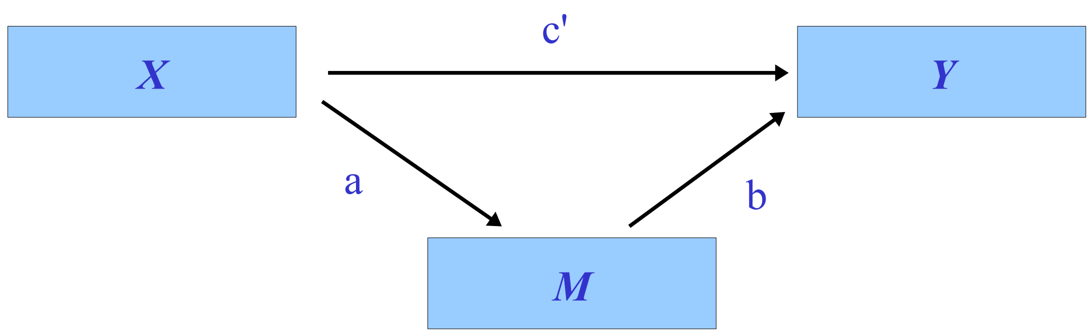
Modello di mediazione

- Il modello di mediazione (semplice) prevede che il processo per cui una variabile X ha un effetto su Y è descrivibile come segue: X ha un effetto su M, M ha un effetto su Y, e perciò X ha un effetto su Y per via dell'intervento di M.

Modello 1



Modello 2



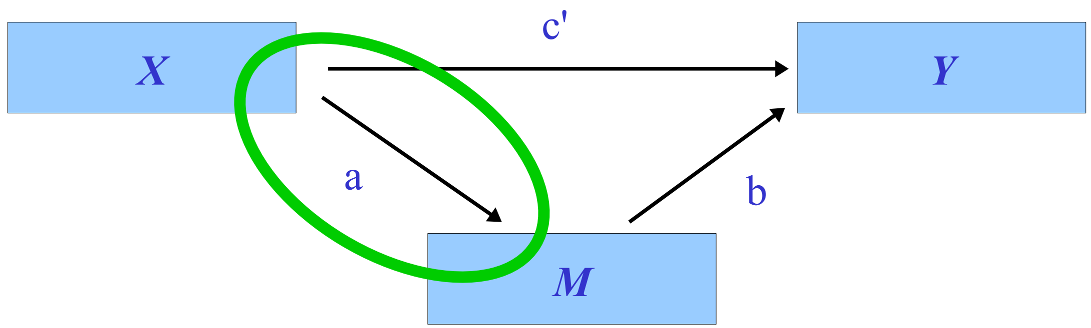
Caratteristiche del mediatore

- Il modello (logico) di mediazione regge se la variabile mediatore possiede alcune caratteristiche:

  - **M deve poter essere causata (o almeno dipendere logicamente) da X**

    Le aspettative devono poter essere influenzate dal prime

Modello 2



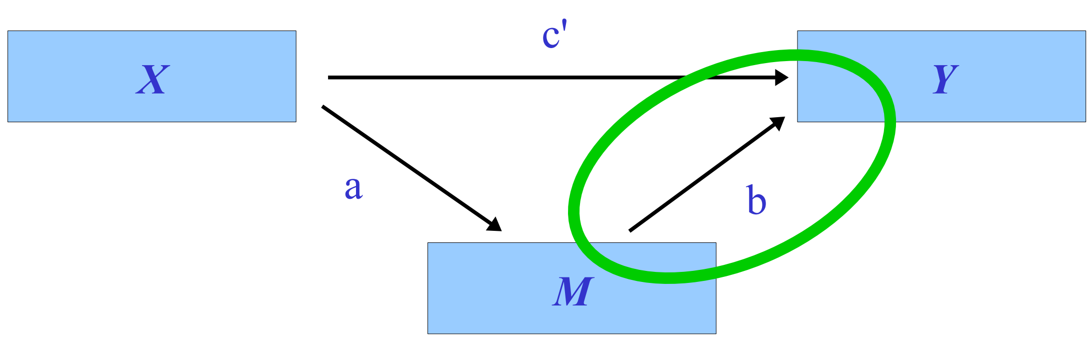
Caratteristiche del mediatore

- Il modello (logico) di mediazione regge se la variabile mediatore possiede alcune caratteristiche:

  - **M deve poter causare (o almeno modificare logicamente) Y**
    Le aspettative devono poter far cambiare il comportamento

  - **M deve poter causare Y indipendentemente da X**
    Le aspettative devono poter far cambiare il comportamento anche in assenza di prime

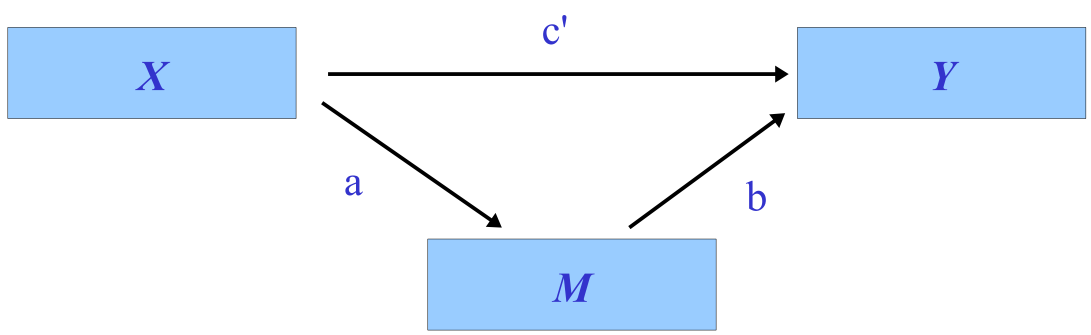
Modello 2



Mediazione

● Se queste caratteristiche sono rispettate (per ora solo logicamente), siamo in presenza di una variabile mediatore, e dunque di un valido modello di mediazione

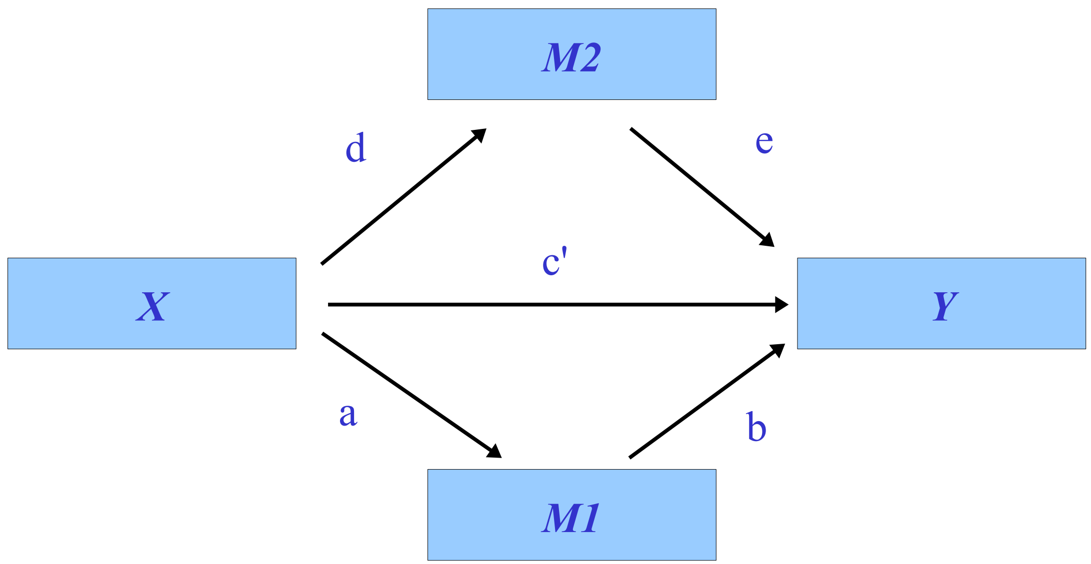
Modello 2



● L'effetto di mediazione sarà quella parte dell'effetto di X su Y che passa per M, cioè che è portato da X ad Y attraverso M

Path analysis

- Il modello logico di mediazione può essere ovviamente esteso a più variabili, dando luogo ad un modello di **path analysis**



Quesito sul “chi”

- Possiamo anche domandarci per chi, o in quali condizioni, PRIME abbia un effetto su BEH

$$\textit{PRIME} \xrightarrow{\text{a}} \textit{BEH}$$

- Possiamo ipotizzare che l'effetto di PRIME non sia uguale per tutti, ma che sia più o meno forte a seconda del tratto di cooperatività

- Ad esempio che l'effetto di PRIME sia più forte se si è cooperativi di proprio, e più debole se si è individualisti.

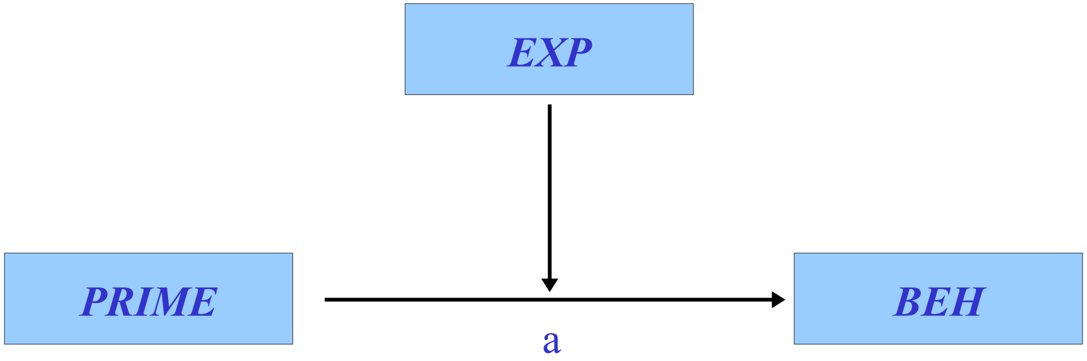
$$\begin{tikzcd}
\textit{PRIME} \arrow[r, shift left, "\text{Forte per Alto SVO}"] \arrow[r, shift right, "\text{Debole per Basso SVO}"'] & \textit{BEH}
\end{tikzcd}$$

Moderazione

- Cioè ipotizziamo che l'effetto di PRIME su EXP non sia uguale per tutti, ma la sua intensità cambi (e.g. cresca) al variare di SVO

- Ipotizziamo che l'effetto di X su Y varia per diversi livelli di M

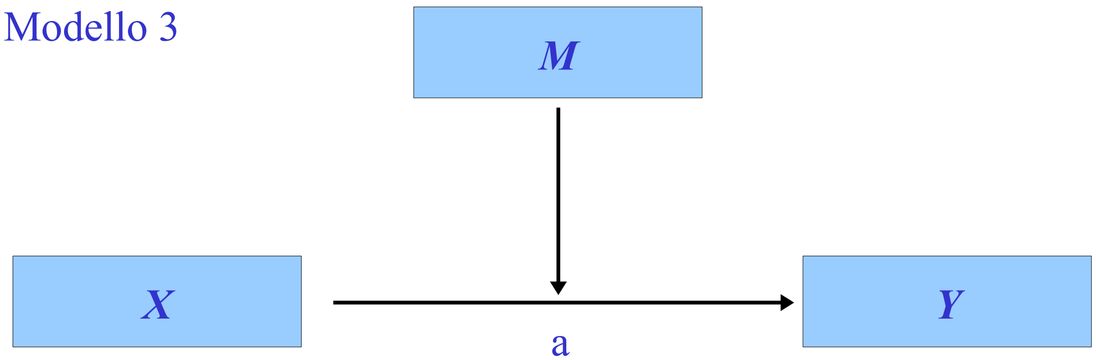
Modello 3



Moderazione

- Se l'intensità dell'effetto di X su Y cambia al variare dei livelli (valori) di un variabile M, diremo che M è un moderatore dell'effetto di X su Y, e che l'effetto di X su Y è condizionale ai valori di M

Modello 3



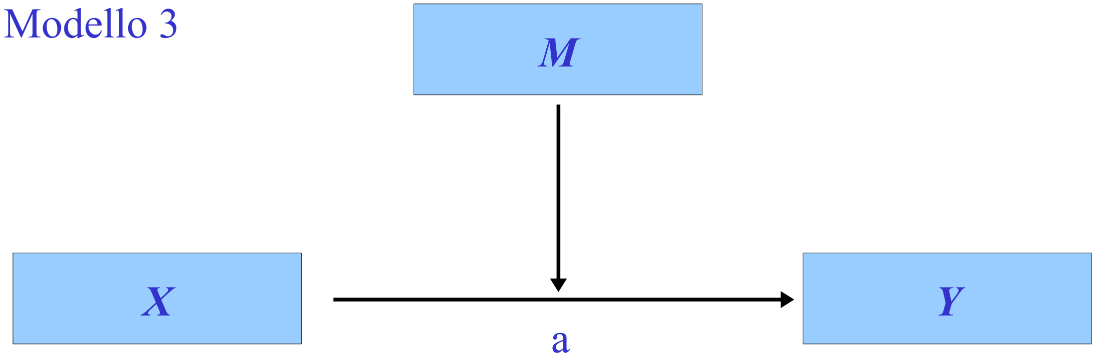
Caratteristiche del moderatore

- Il modello (logico) di moderazione regge se la variabile moderatore possiede alcune caratteristiche:

  - **M deve poter cambiare l'intensità dell'effetto tra X e Y**
    SVO descrive persone differenti che possono essere più o meno sensibili al PRIME

  - **M non è generalmente causato da X**
    SVO è un tratto e non dipende dal prime ricevuto

Modello 3



Modello Multiplo vs Mediazione vs Moderazione

- I tre modelli teorici sono molto differenti e rispondono a domande diverse

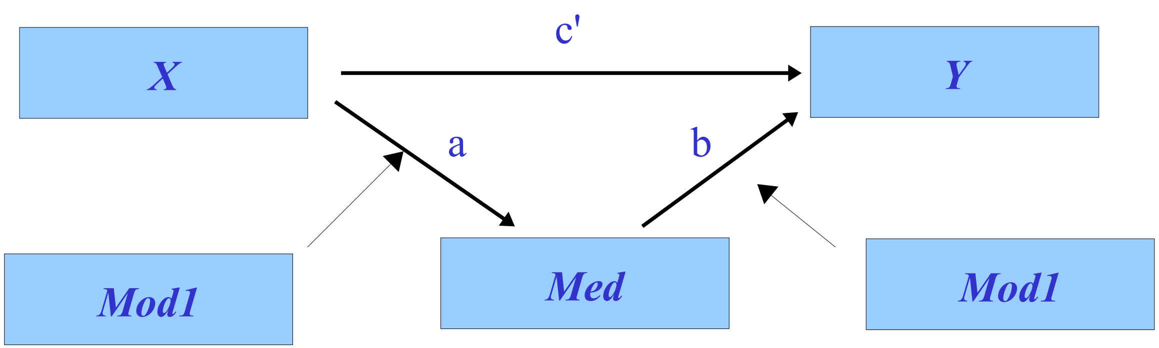
Covariata	Mediatore	Moderatore
Risponde alla domanda: "al netto di ..." I predittori sono tutti allo stesso livello Ogni effetto è calcolato covariando gli altri	Risponde alla domanda: "perchè" Può essere causato da X Non modifica l'effetto, lo assorbe e lo trasferisce	Risponde alla domanda: "chi", "in quali condizioni" E' indipendente da X Modifica l'effetto

Mediazione e Moderazione

- I modelli teorici **possono operare insieme per spiegare gli effetti**

Mediazione condizionale o moderata

Modello 4



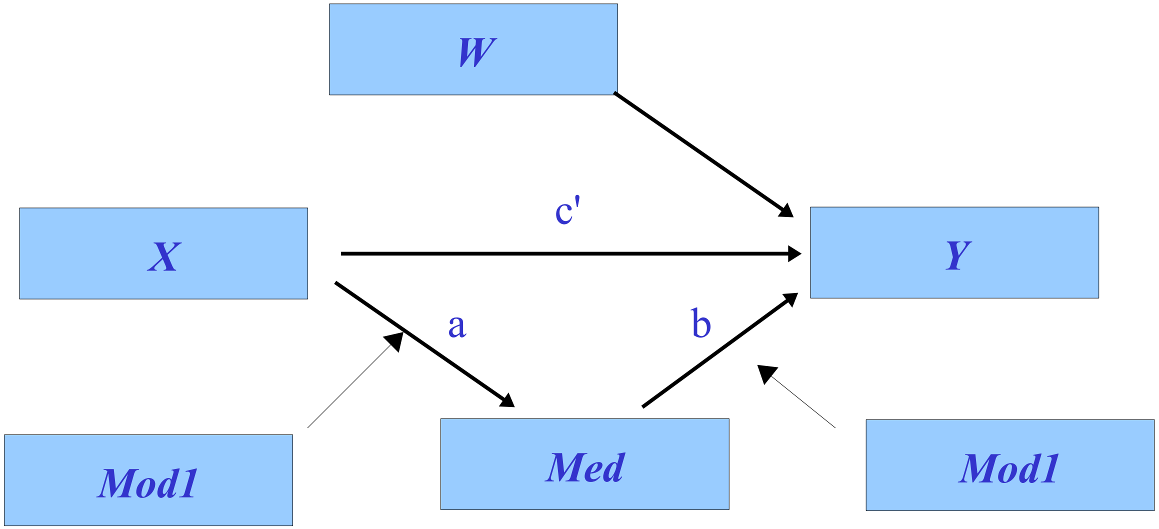
- Cioè alcuni effetti inerenti al modello di mediazione sono condizionali ai livelli di un'altra variabile moderatore. Il modello di mediazione cambia per diversi livelli del moderatore (i)

Mediazione e Moderazione

- I modelli teorici **possono operare insieme per spiegare gli effetti**

Mediazione condizionale o moderata con covariate

Modello 5



- Cioè alcuni effetti inerenti al modello di mediazione sono condizionali ai livelli di un'altra variabile moderatore. Il modello di mediazione cambia per diversi livelli del moderatore (i)

Applicazione ai Modelli Statistici Generali

Applicazione

- I modelli multipli, di mediazione e moderazione si possono stimare statisticamente mediante qualunque tipo di modello statistico che quantifichi la relazione tra variabili mediante un coefficiente numerico

- Quale modello statistico utilizzare dipende dal tipo di variabili del disegno di ricerca (continue, categoriche) ed il tipo di disegno di ricerca (cross-sectional, longitudinale, a misure ripetute)

- Il modello lineare generale (*regressione/anova*)
- Il modello lineare misto (*random coefficients models*)
- Il modello lineare generalizzato (*logistica, poisson, multinomiale*)

GLM

- La maggior parte delle tecniche più usate (almeno in psicologia), come Regressione, correlazione, ANOVA, sottostanno ad un unico modello lineare generale

Modello Lineare Generale

$$y_i = a + b_1 \cdot x_{1i} + b_2 \cdot x_{2i} + \, ..b_k \cdot x_{ki} + e_i$$

Variabile dipendente

Variabili indipendenti

Errore

GLM

Modello Lineare Generale

vantaggi

- Consente di stimare le relazioni fra due o più variabili
- Si applica ad una ampio spettro di tipi di dati
- Consente di stimare vari tipi di effetti

svantaggi

- Assume una struttura dei dati molto semplice
- Non consente di modellare una ampia serie di relazioni e dipendenza tra unità di misurazione

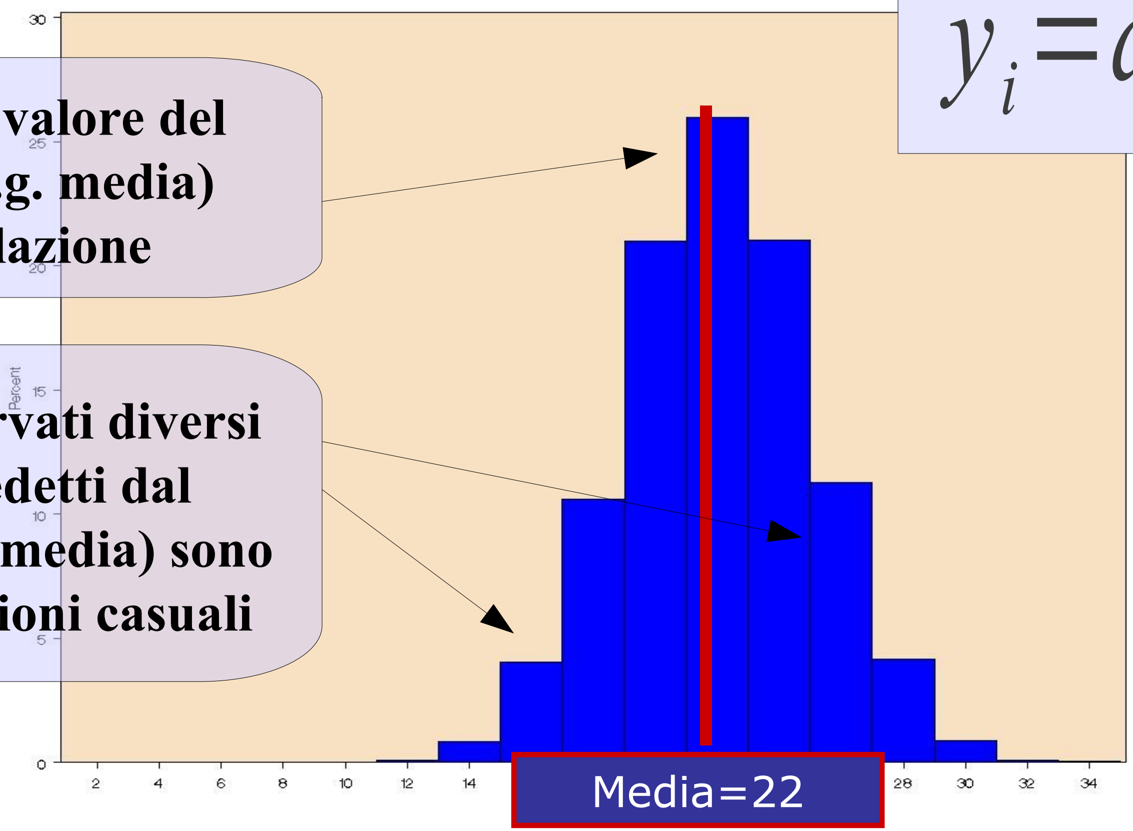
Assunzioni GLM

Modello Lineare Generale

$$y_i = a + e_i$$

Esiste un solo valore del parametro (e.g. media) nella popolazione

I punteggi osservati diversi da quelli predetti dal parametro (e.g. media) sono dovuti a variazioni casuali



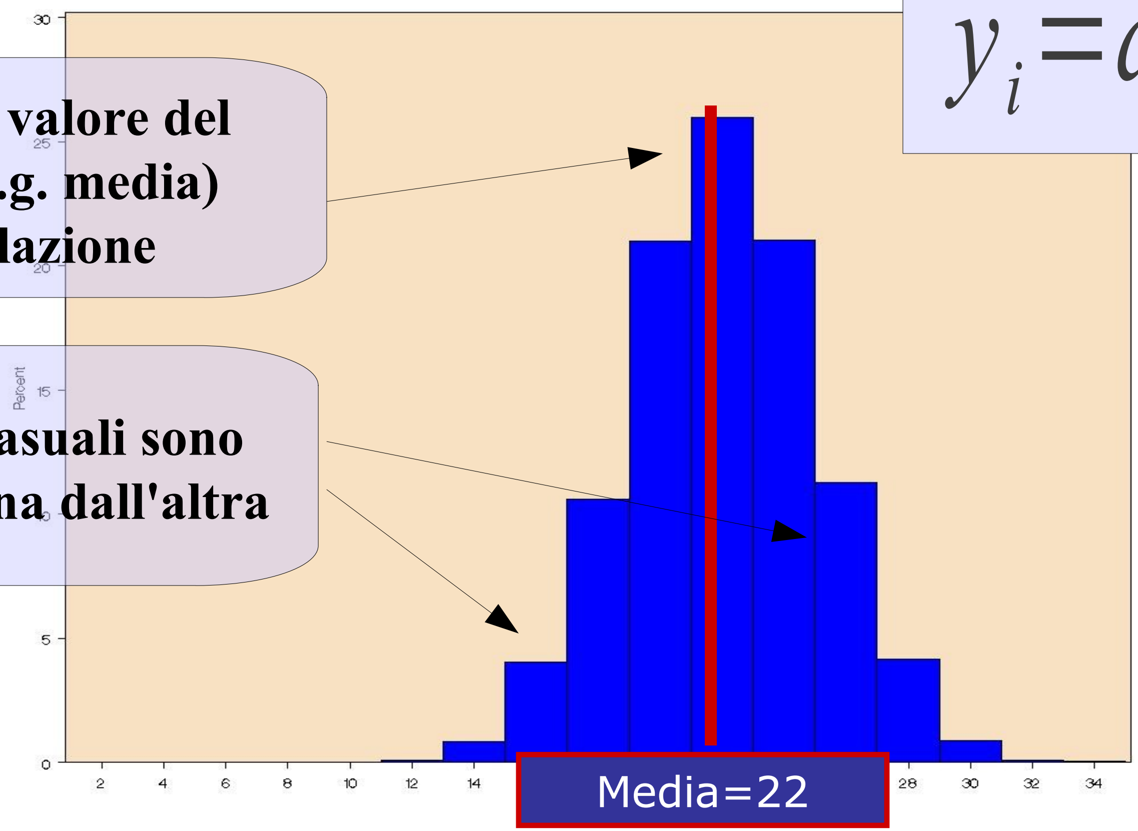
Assunzioni GLM

Modello Lineare Generale

$$y_i = a + e_i$$

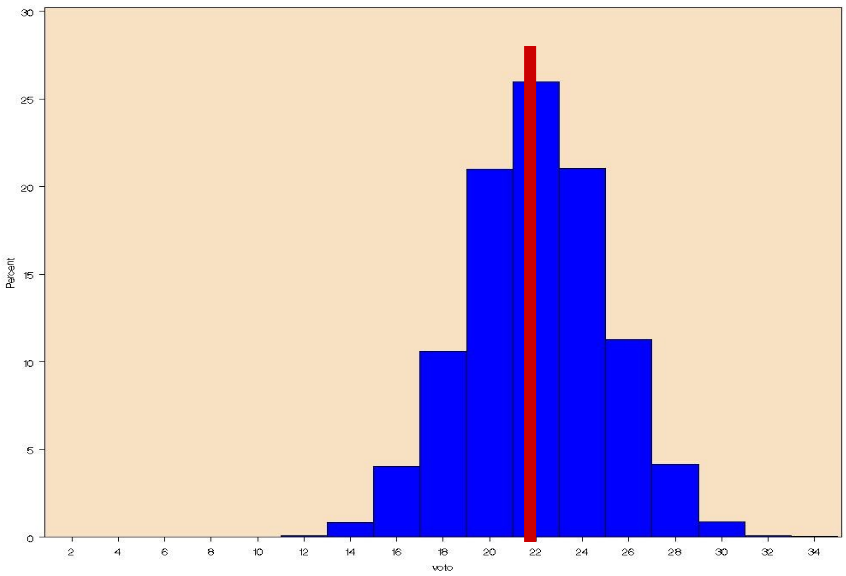
Esiste un solo valore del parametro (e.g. media) nella popolazione

Le variazioni casuali sono indipendenti l'una dall'altra



Assunzioni GLM

Modello Lineare Generale



Il valore stimato della popolazione si definice FISSO (fixed parameter)

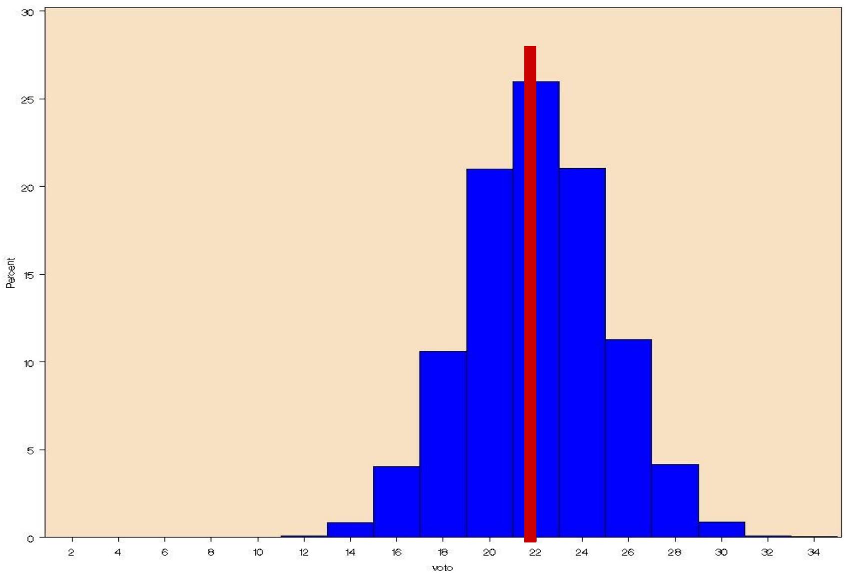
$$y_i = a + e_i$$

$$corr(e_i, e_j) = 0$$

Le variazioni casuali sono indipendenti l'una dall'altra

Assunzioni GLM

Modello Lineare Generale



$$y_i = a + e_i$$

$$e_i \quad N(0, \sigma^2)$$

I residui del modello sono distribuiti normalmente

Generalizzazioni

- Useremo il Modello Lineare Misto (*random coefficients models*) quando le assunzioni di unicità degli effetti e l'indipendenza dei residui non è rispettata (misure ripetute, dati clusterizzati)

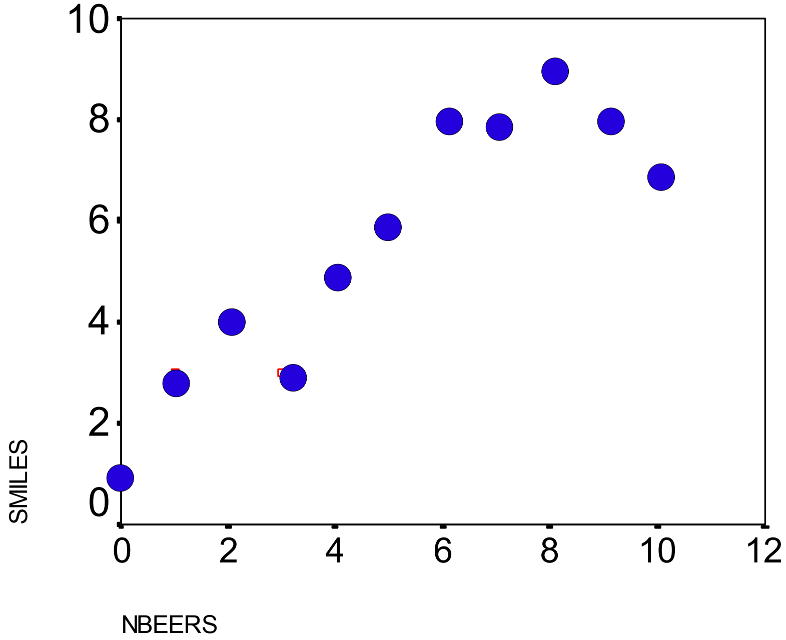
- Useremo il Modello Lineare Generalizzato quando le assunzioni di normalità dei residui non può essere rispettata (variabili dipendenti categoriche)

Il modello di regressione

Concetti fondamentali

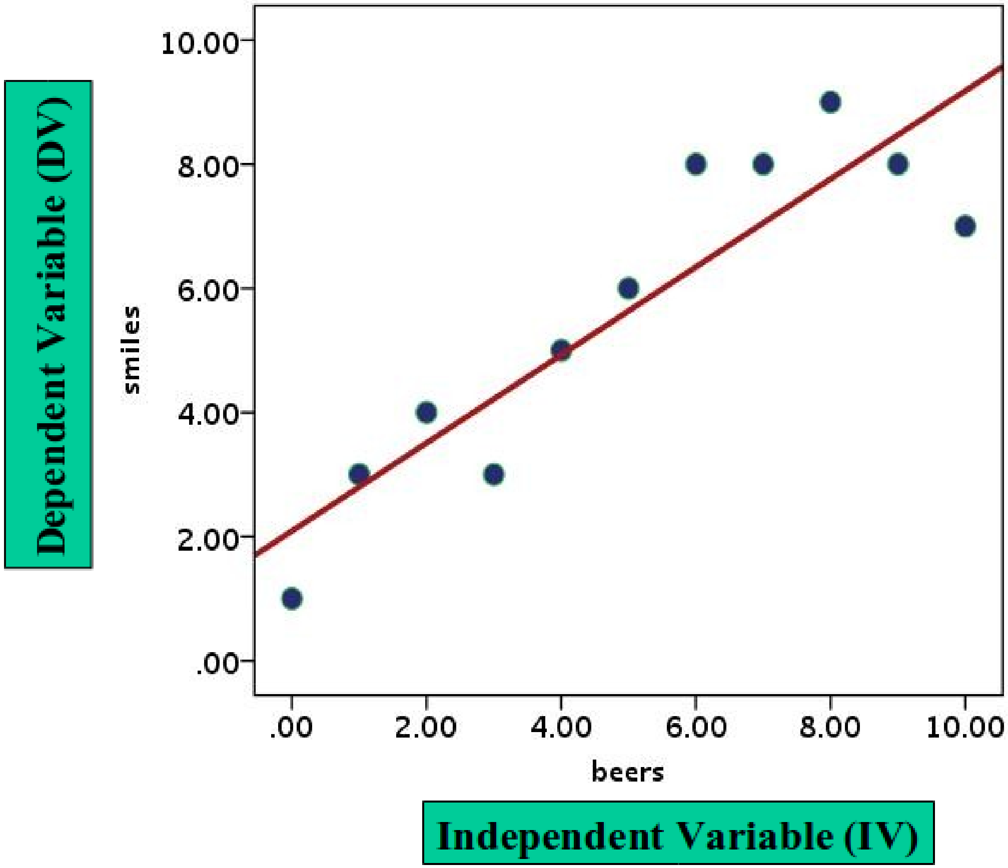
- Consideriamo ora questa ipotetica ricerca: siamo andati in un pub ed abbiamo contato quanti sorrisi le persone ai tavoli producevano (ogni 10 minuti) e quante birre avevano bevuto fino a quel momento

Birre	Sorrisi
0	1
1	3
2	4
3	3
4	5
5	6
6	8
7	8
8	9
9	8
10	7



Concetti fondamentali

- Lo scopo della retta di regressione è di rappresentare la relazione lineare tra la variabile indipendente e la dipendente



Nel caso più semplice, abbiamo una retta semplice

$$\hat{y}_i = a + b \cdot x_i$$

Concetti fondamentali

La retta può essere descritta mediante due coefficienti: il termine costante ed il coefficiente angolare (slope)

Termine constante
(o intercetta)

$$\hat{y}_i = a + b \cdot x_i$$

Coefficiente B
(o slope)

Concetti fondamentali

Nel nostro esempio...usando SPSS

Termine constante
(o intercetta)

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	2.091	.684		3.057	.014
	NBEERS	.709	.116	.898	6.132	.000

a. Dependent Variable: SMILES

$$\hat{y}_i = a + b \cdot x_i$$

Coefficiente B
(o slope)

Concetti fondamentali

Nel nostro esempio... in R

Termine constante
(o intercetta)

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.0909     0.6841   3.057 0.013647 *
## beers         0.7091     0.1156   6.132 0.000172 ***
## ---
```

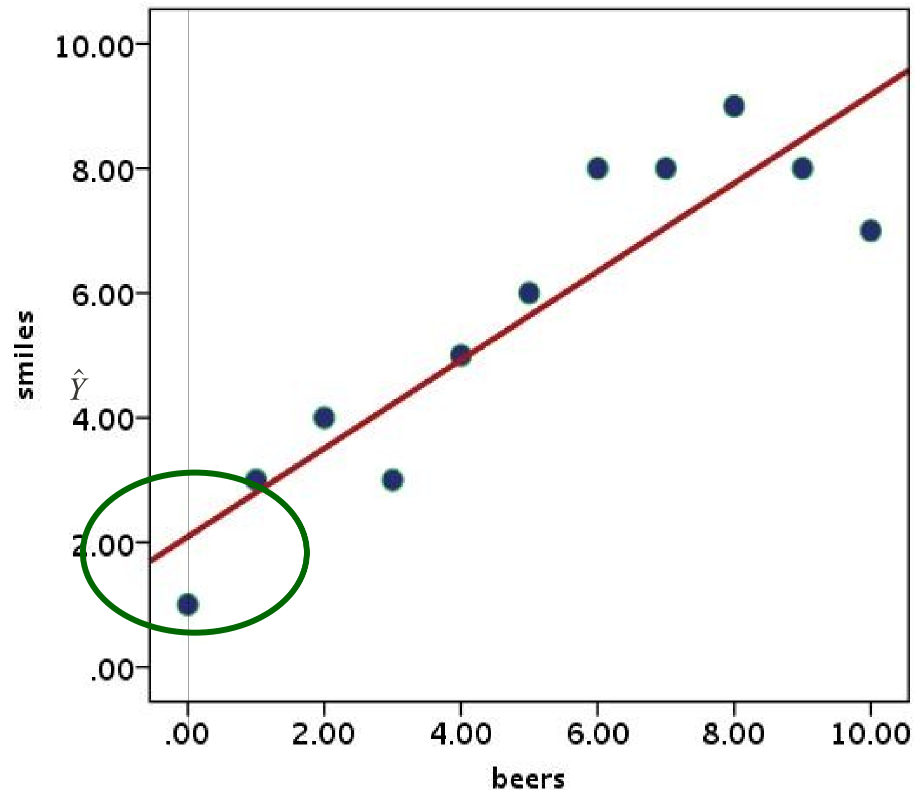
$$\hat{y}_i = a + b \cdot x_i$$

Coefficiente B
(o slope)

Costante o intercetta

a l'intercetta della retta: indica il valore atteso (medio) della VD per la VI=0

$$\hat{y} = a + b \cdot 0$$

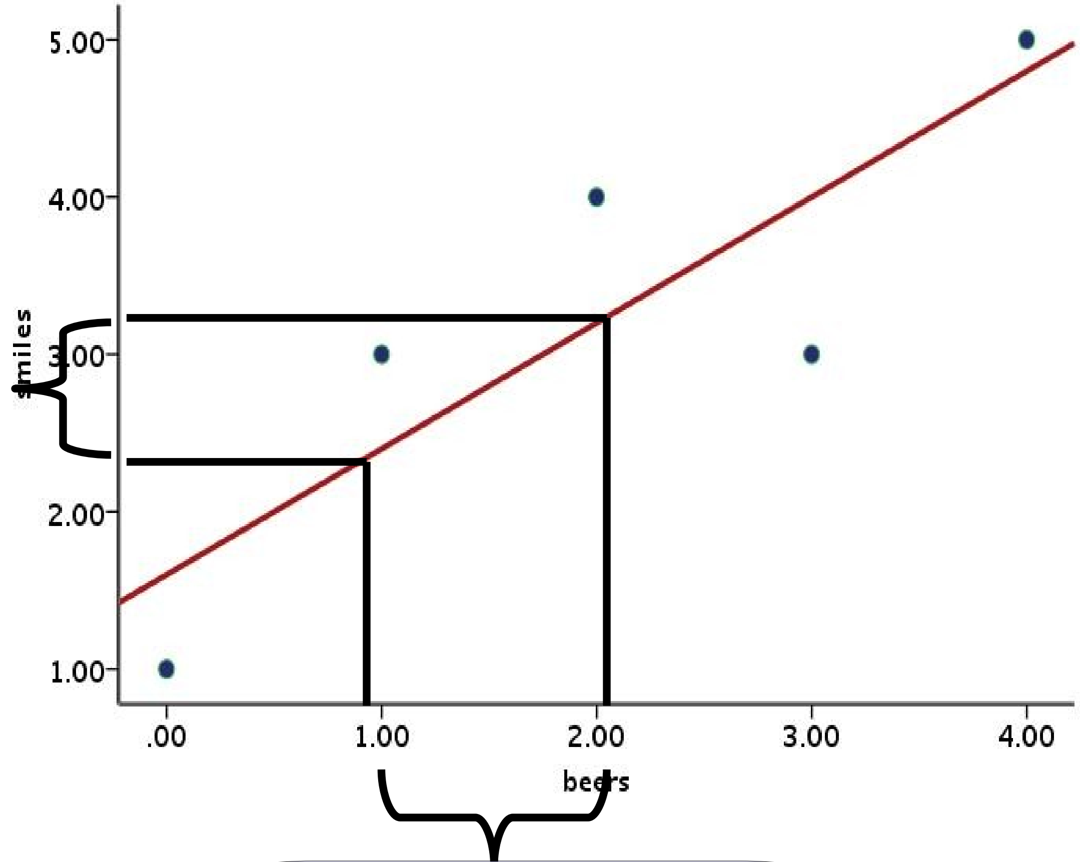


Quando un partecipante ha bevuto zero birre, mostra (in media) 2.09 sorrisi

Coefficiente di regressione

B è il coefficiente angolare della retta: indica il cambiamento atteso nella VD al variare di una unità della VI

I sorrisi aumentano di B unità



Per una unità in più della VI: una birra in più

Per ogni birra che si beve, i sorrisi aumentano in media di .709 unità

Test inferenziale

I coefficenti vengono testati per la loro significatività statistica mediante il t-test **t test**

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.0909     0.6841   3.057 0.013647 *
## beers         0.7091     0.1156   6.132 0.000172 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Se Pr. < 0.05, diremo che B è significativamente diverso da zero

Soluzione standardizzata

- A volte conviene ottenere un coefficiente che non dipenda dalle unità di misura, cioè un coefficiente standardizzato (beta)

- In qualunque modello lineare il coefficiente standardizzato si ottiene stimando il modello sulle variabili standardizzate (z-score)

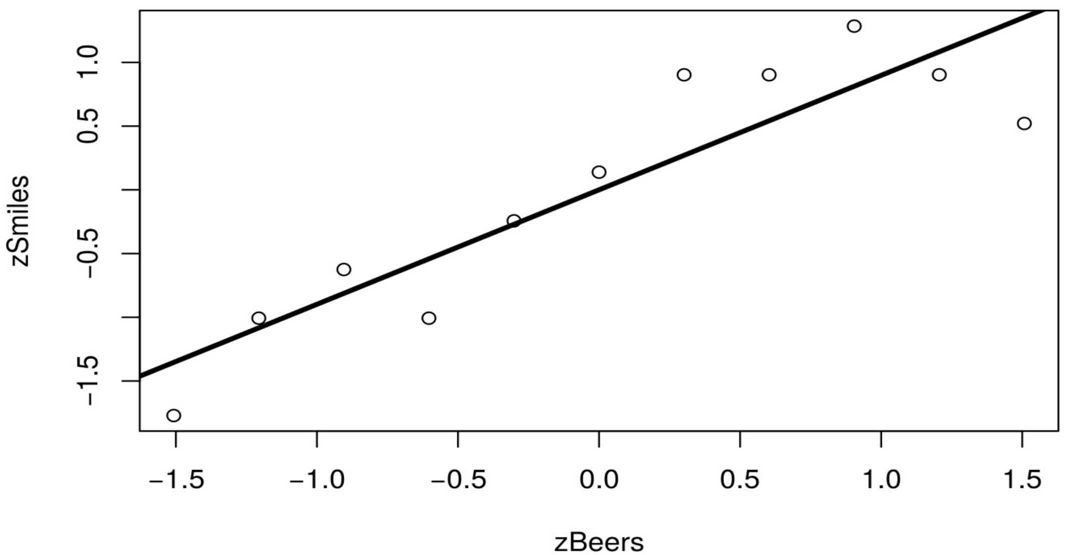
  - Prima si standardizzano le variabili

  - Poi si ricalcola il modello

Soluzione standardizzata

- Il beta (nella regressione singola) altro non è che il coefficiente di correlazione r di pearson

```
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.000      0.140    0.00  1.00000
## zbeers           0.898      0.146    6.13  0.00017 ***
## ---
```



R

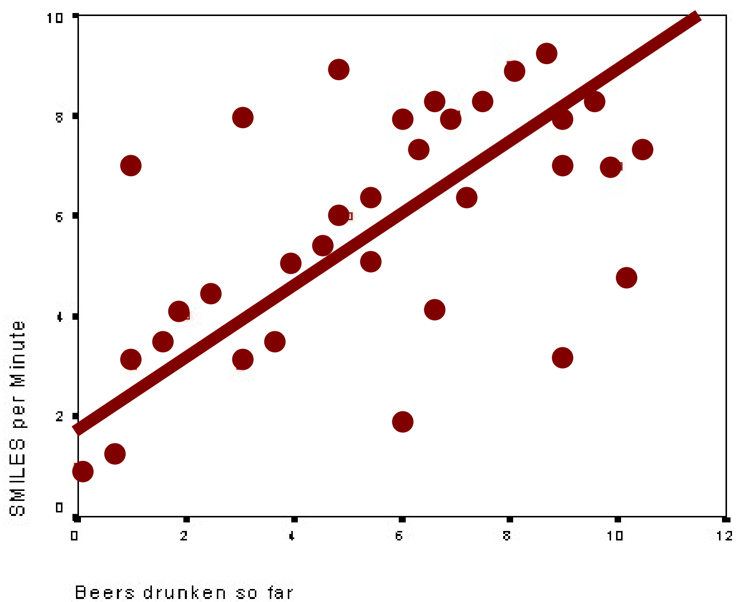
- Per tutti i modelli inerenti al modello lineare generale, la funzione R è sempre la stessa, assai semplice da ricordare

```r
#stimo la regressione
mod<-lm(smiles~beers,data=dat)
```

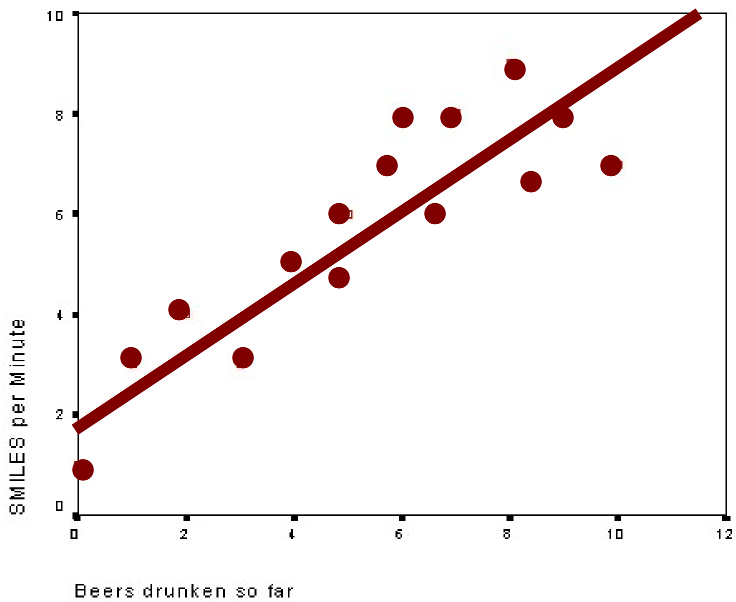
Bontà di adattamento

Non tutte le rette di regressione hanno lo stesso potere predittivo, cioè la stessa capacità di adattarsi ai dati osservati

bassa



alta



Errore di regressione

Notiamo che la predizione non corrisponde di norma ai valori osservati

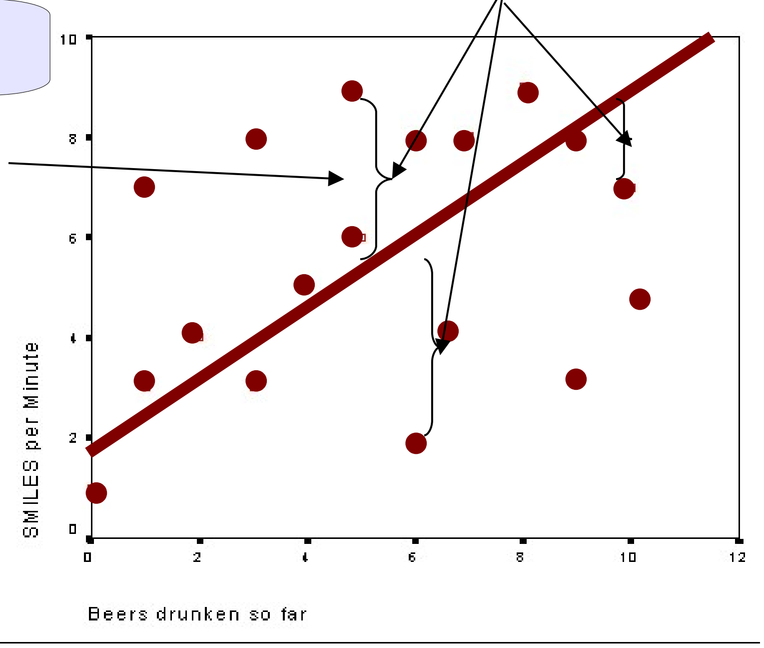
$$\hat{y}_i = a + b_{yx} x_i$$

Predetti

Discrepanza

Errore

$$y_i - \hat{y}_i = y_i - (a + b_{yx} x_i)$$



I valore di Y sono decomposti in predetti e residui

$$y_i = \underbrace{(a + b_{yx} x_i)}_{\text{retta}} + \underbrace{(y_i - \hat{y}_i)}_{\text{residui}}$$

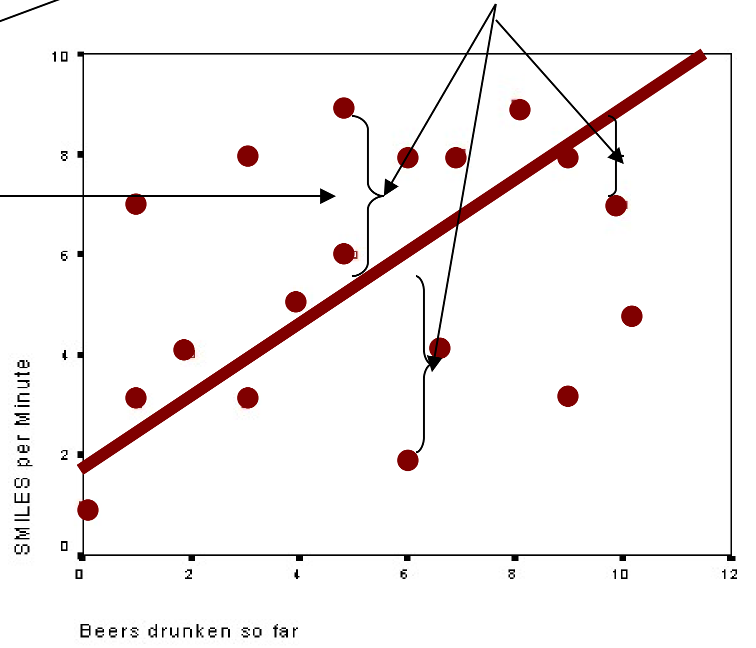
Quanto e' grande l'errore di regressione

Calcoliamo la distanza media tra i punti osservati e la retta

Le distanze si calcolano mediante le discrepanza al quadrato

discrepanze osservati-predetti

$$\sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n-1} = s_e^2$$



Notiamo che questa è una varianza che chiameremo varianza di errore

Proporzione riduzione errore

Il modello si adatterà ai dati tanto più riduce l'errore di predizione rispetto a non usare tale modello

- La logica è di confrontare due casi:
  - L'errore calcolato per la regressione data
  - L'errore associato alla media, cioè errore associato a non utilizzare la regressione

Proporzione riduzione errore

Senza regressione l'unica predizione plausibile di Y e' la media di Y

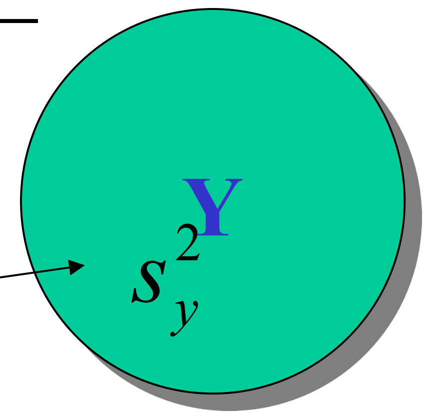
Predizione senza regressione

$$\hat{y}_i = M_y$$

Varianza di errore senza predizione

$$s_y^2 = \frac{\sum (y_i - M_y)^2}{n-1}$$

Le deviazioni dalla media (la varianza) non siamo in grado di spiegarle



Proporzione riduzione errore

- Con la regressione faremo una certa predizione

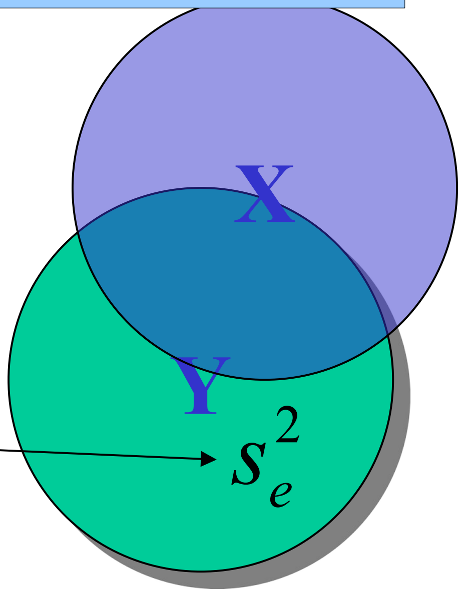
Predizione con regressione

$$\hat{y}_i = a + b_{yx} x_i$$

Varianza di errore con predizione

$$s_e^2 = \frac{\sum (y_i - \hat{y}_i)^2}{n-1}$$

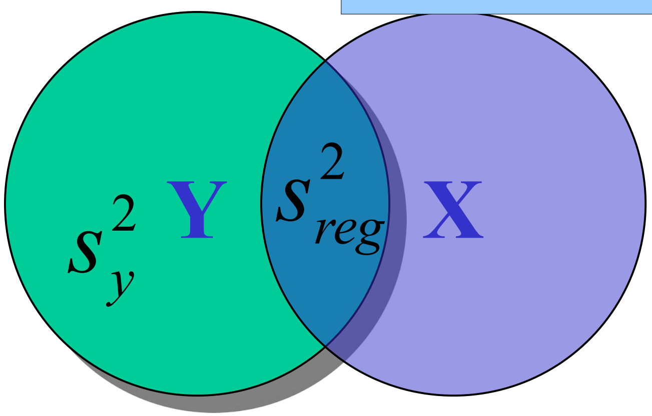
Le deviazioni dalla regressione (varianza di errore) non siamo in grado di spiegarle



R-quadro

- Dunque il fit della regressione è tanto buono quanto riesce a migliorare la predizione, cioè a diminuire l'errore

Errore senza regressione



$$\frac{s_y^2 - s_e^2}{s_y^2} = \frac{s_{reg}^2}{s_y^2} = R_{yx}^2$$

Errore con regressione

Il coefficiente R2 indica la proporzione di errore ridotto dalla regressione, anche detta **Varianza Spiegata**

R-quadro

- Nel nostro esempio...in SPSS

Tests of Between-Subjects Effects

Dependent Variable: smiles

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	55.309[a]	1	55.309	37.607	.000
Intercept	13.740	1	13.740	9.343	.014
beers	55.309	1	55.309	37.607	.000
Error	13.236	9	1.471		
Total	418.000	11			
Corrected Total	68.545	10			

a. R Squared = .807 (Adjusted R Squared = .785)

- L'ipotesi nulla che R^2 sia zero viene testata con il F-Test

R-quadro

- Nel nostro esempio...in R

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.0909     0.6841   3.057 0.013647 *
## beers         0.7091     0.1156   6.132 0.000172 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.213 on 9 degrees of freedom
## Multiple R-squared:  0.8069, Adjusted R-squared:  0.7854
## F-statistic: 37.61 on 1 and 9 DF,  p-value: 0.0001723
```

- L'ipotesi nulla che R^2 sia zero viene testata con il F-Test

Output di un modello lineare

- Ogni modello lineare può essere interpretato sia in termini di varianze

Tests of Between-Subjects Effects

Dependent Variable: smiles

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	55.309[a]	1	55.309	37.607	.000
Intercept	13.740	1	13.740	9.343	.014
beers	55.309	1	55.309	37.607	.000
Error	13.236	9	1.471		
Total	418.000	11			
Corrected Total	68.545	10			

a. R Squared = .807 (Adjusted R Squared = .785)

- Che in termini di coefficienti

Coefficients[a]

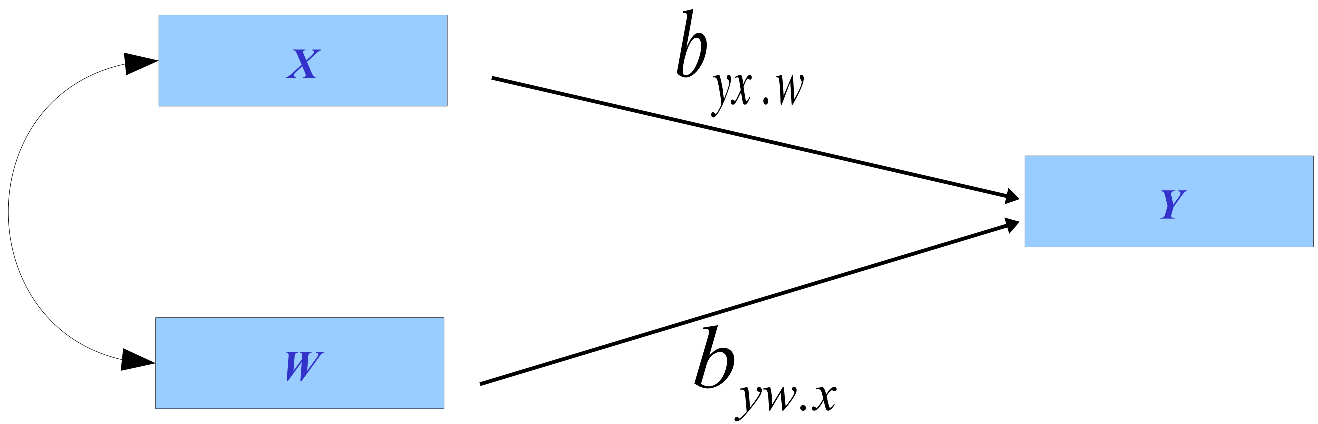
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.091	.684		3.057	.014
	NBEERS	.709	.116	.898	6.132	.000

a. Dependent Variable: SMILES

Effetti multipli

- Consideriamo ora il caso in cui la variabile dipendente possa essere spiegata da più di una variabile

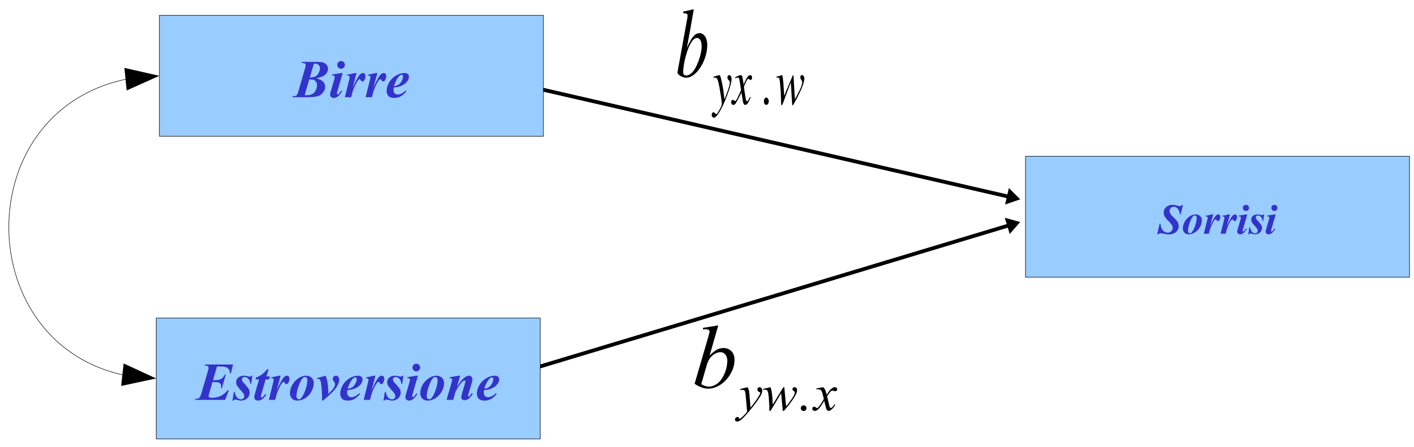
Regressione Multipla



Esempio Effetti multipli

- Vogliamo predire *il numero di sorrisi* sia con *il numero di birre* che con *il tratto "estroversione"* del soggetto

Regressione Multipla

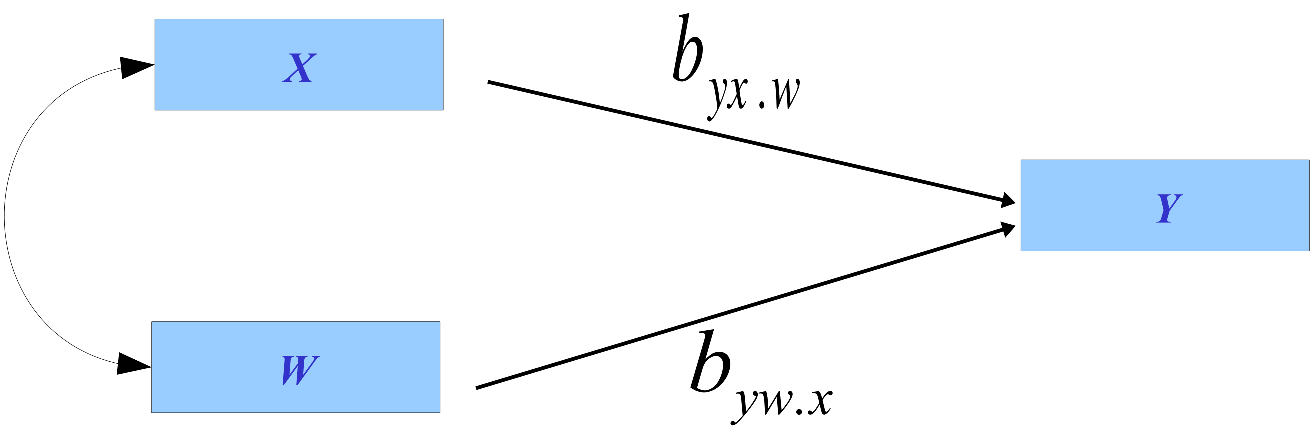


Effetti multipli

- La regressione aggiunge termini lineari per ogni variabile indipendente

Regressione Multipla

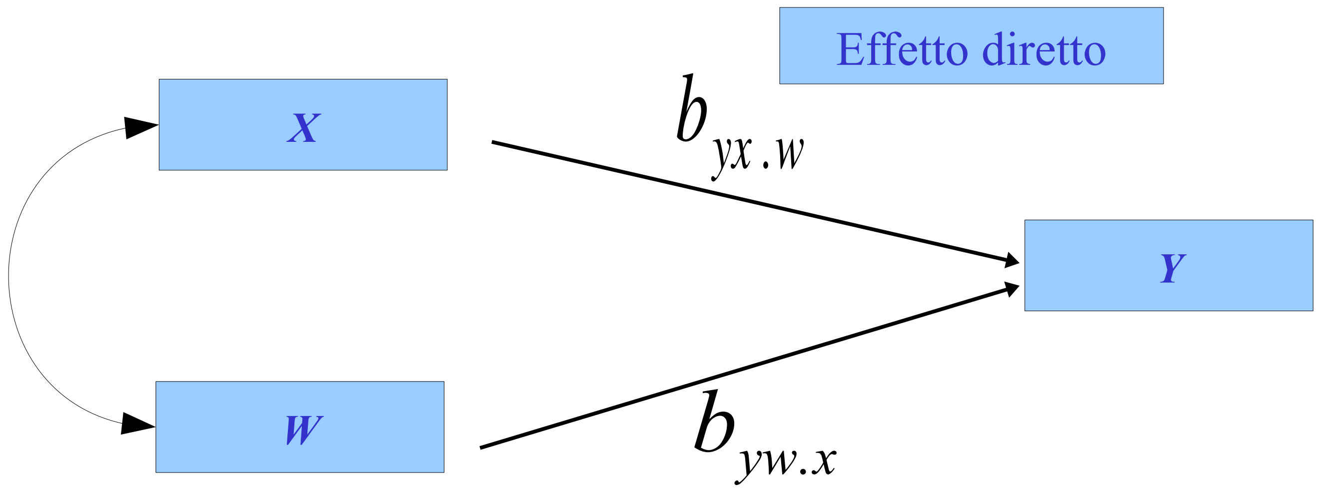
$$\hat{y} = a + b_{yx.w}x + b_{yw.x}w$$



Effetti multipli

- Ogni coefficiente di regressione esprime l’effetto diretto della IV su Y, togliendo l’effetto che passa indirettamente per l'altra VI

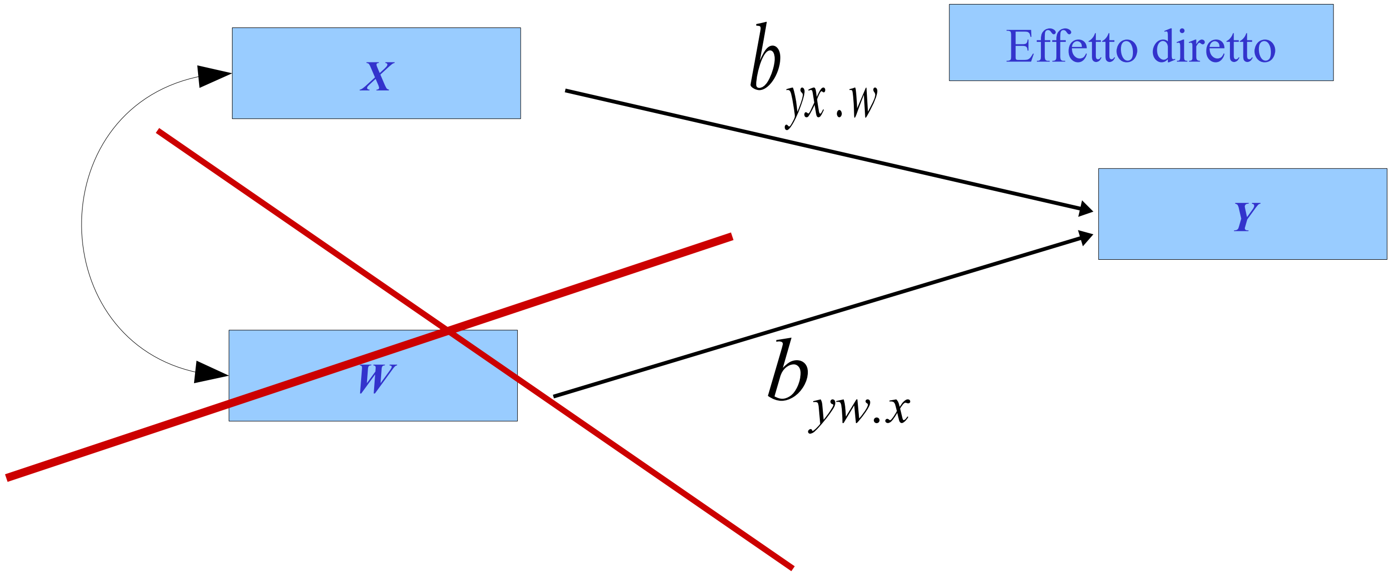
Effetti parziali



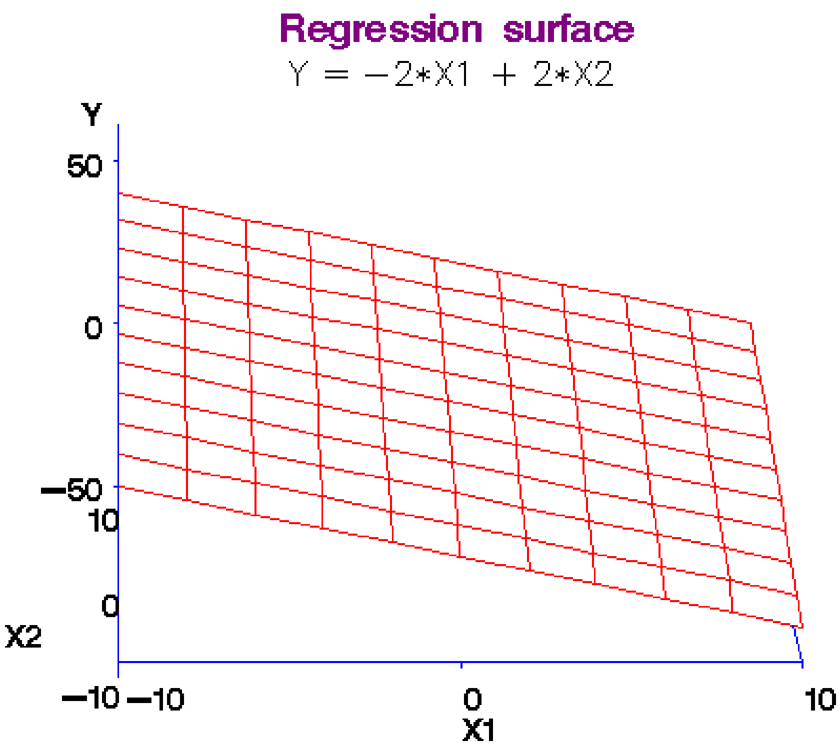
Effetti multipli

- Togliere l’effetto indiretto è equivalente a bloccare la possibilità che x vada su y mediante w: Il coefficiente viene dunque detto **coefficiente parziale, cioè l’effetto di x parzializzando l’effetto di w**

Effetti parziali

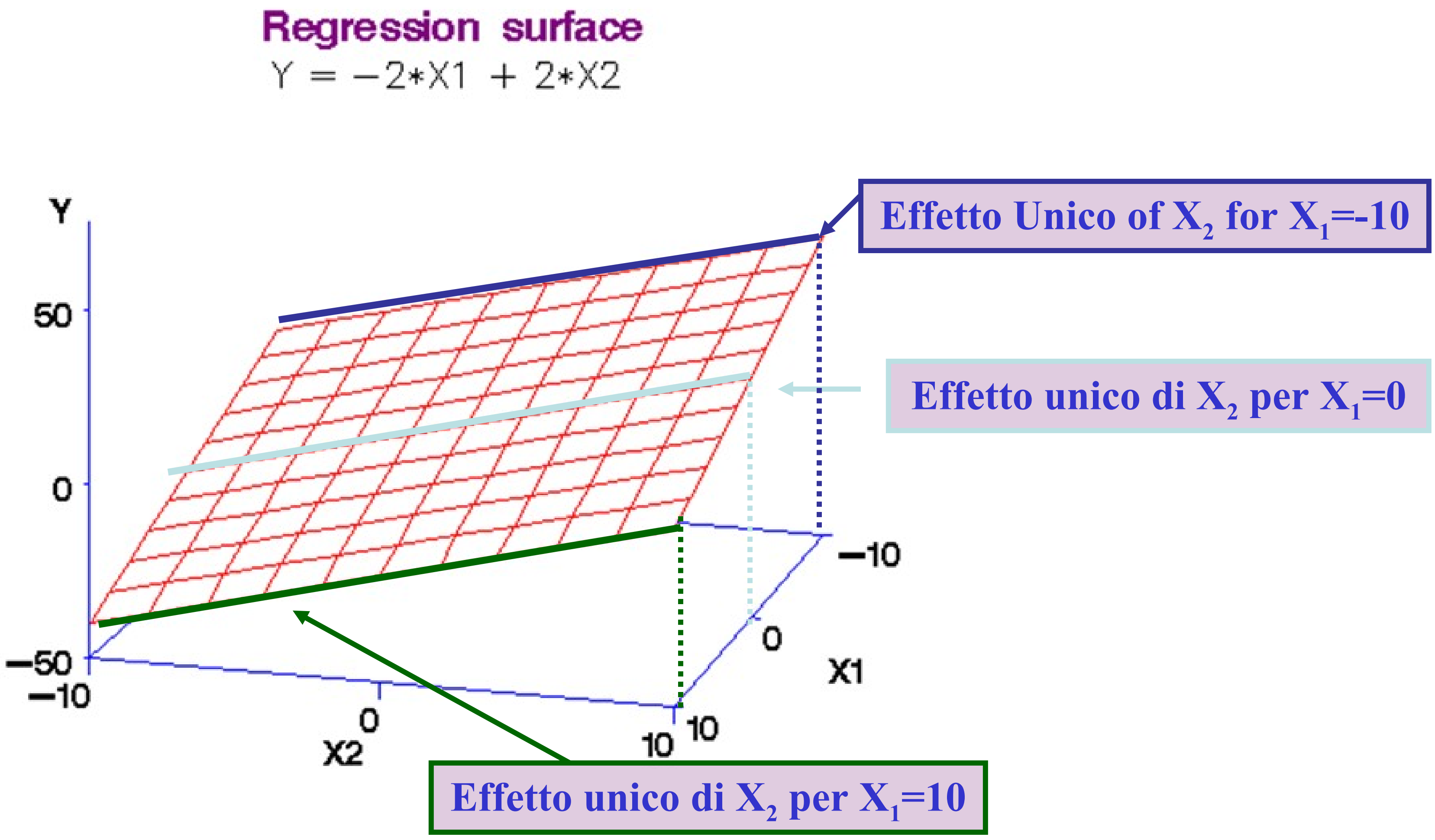


Rappresentazione geometrica



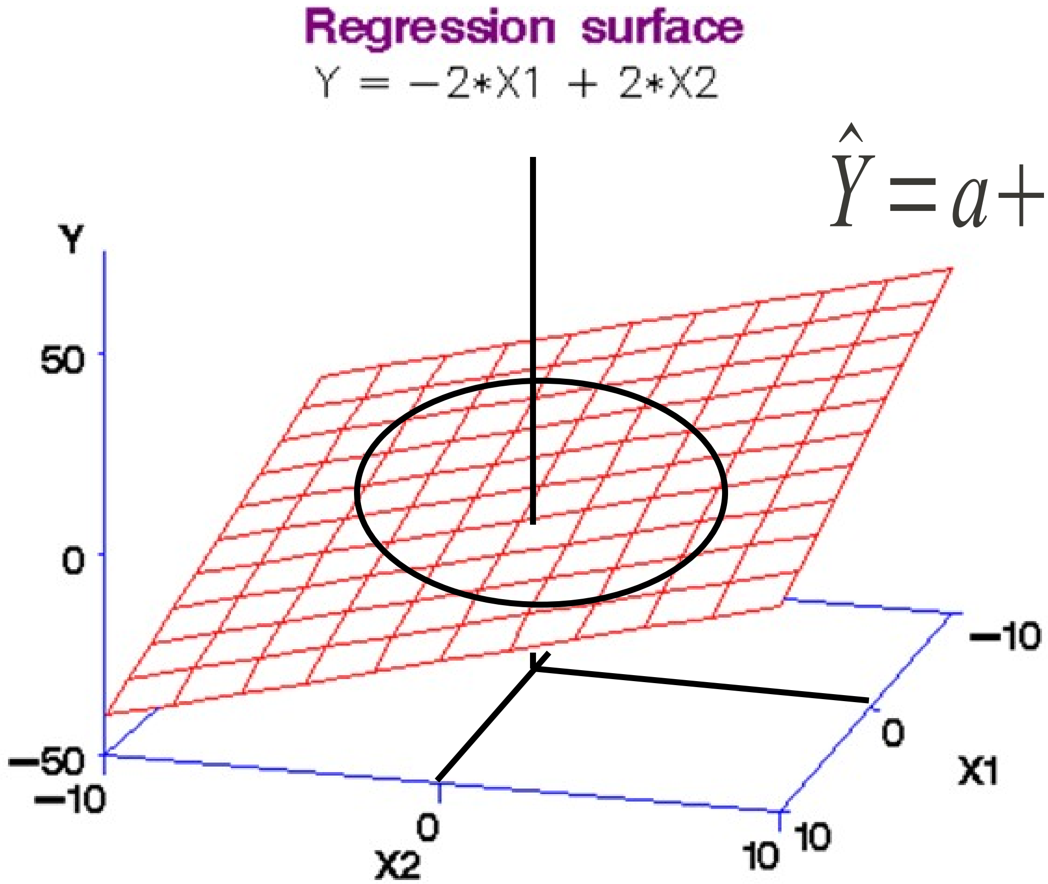
$$\hat{y} = a + B_{y1}x_1 + B_{y2}x_2$$

Interpretazione geometrica



Intercetta (o costante)

- L'intercetta indica il valore atteso della VD per tutte le VI uguali a 0



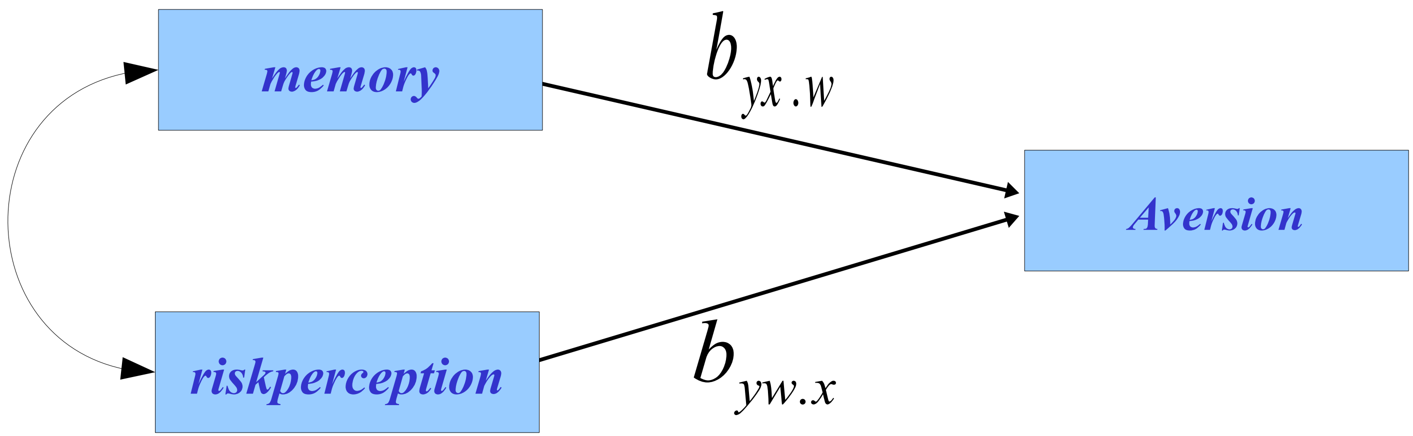
$$\hat{Y} = a + B_{y1.2}0 + B_{y2.1}0$$

$$\hat{Y} = a$$

Esempio

- Una campagna pubblicitaria contro il fumo è stata testata chiedendo ai partecipanti di ricordare il maggior numero di spot della campagna (misura di esposizione) (*memory*), i rischi percepiti del fumo (*riskperception*), e l'avversione al fumo (*avversion*).

- Supponiamo di voler vedere se l'esposizione alla campagna abbia un effetto sull'avversione, considerando anche i rischi percepiti.



Esempio

- Effetti parziali delle VI sulle VD

```
## lm(formula = aversion ~ memory + riskperception, data = smoke)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -64.489   -6.869    1.276    8.542   38.694
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -73.66753    6.57749 -11.200   <2e-16 ***
## memory          1.97548    1.90592   1.036    0.303
## riskperception  1.44118    0.08558  16.839   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16.67 on 97 degrees of freedom
## Multiple R-squared:  0.7631, Adjusted R-squared:  0.7582
## F-statistic: 156.2 on 2 and 97 DF,  p-value: < 2.2e-16
```

Coefficienti B

t.test e p

Esempio

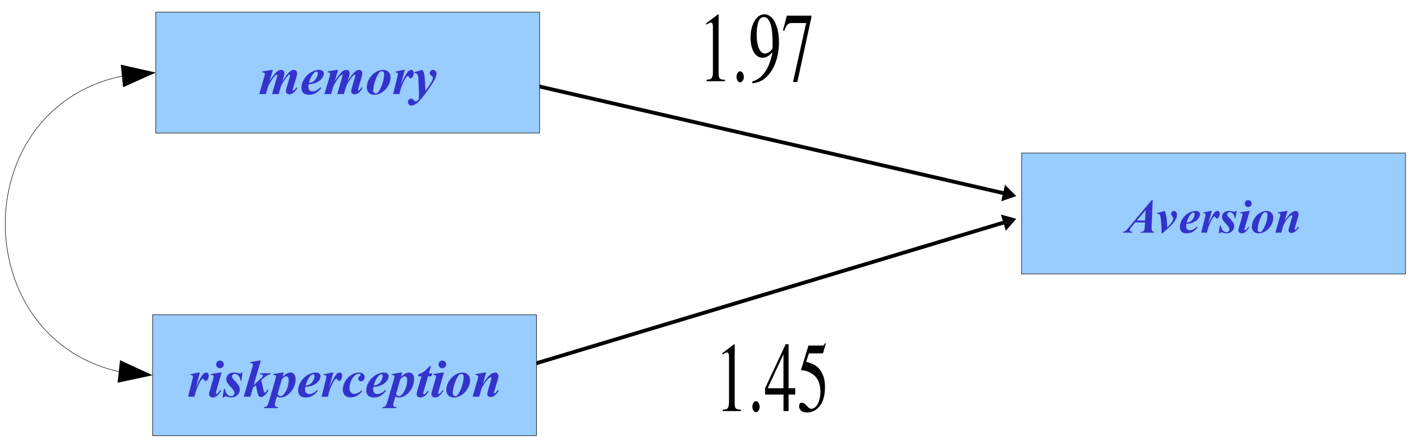
- Effetti parziali delle VI sulle VD

```
## lm(formula = aversion ~ memory + riskperception, data = smoke)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -64.489   -6.869    1.276    8.542   38.694
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -73.66753    6.57749 -11.200   <2e-16 ***
## memory            1.97548    1.90592   1.036    0.303
## riskperception    1.44118    0.08558  16.839   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16.67 on 97 degrees of freedom
## Multiple R-squared:  0.7631, Adjusted R-squared:  0.7582
## F-statistic: 156.2 on 2 and 97 DF,  p-value: < 2.2e-16
```

- Al netto di memory, riskperception ha un effetto di B=1.44, t(97)=16.83, p.<.001 mentre al netto di riskperception, memory non ha un effetto, B=1.97, t(97)=1.036, p.=.303

Esempio

- Modello finale



Esempio

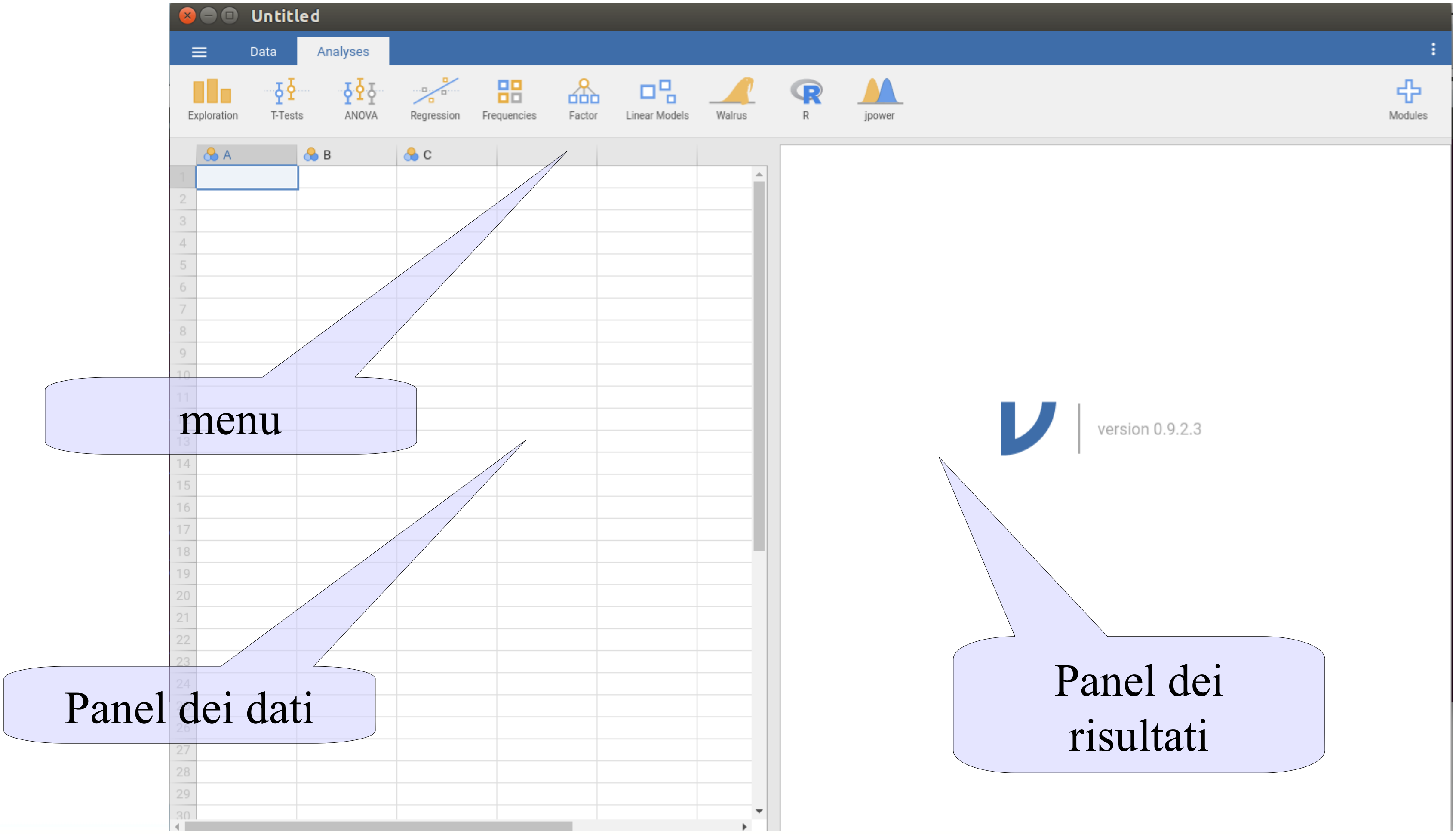
- Facendo una regressione semplice tra memory e aversion, i risultati sono differenti

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -25.943     11.697  -2.218  0.02887 *
## memory         9.933      3.639   2.730  0.00751 **
## ---
```

- L'effetto di *memory* (alto e significativo) si riduce a zero quando *riskperception* è tenuto costante. In altre parole, se tutti avessero lo stesso livello di *riskperception*, il ricordo della campagna non avrebbe effetto sull'avversione (possibile mediazione?)

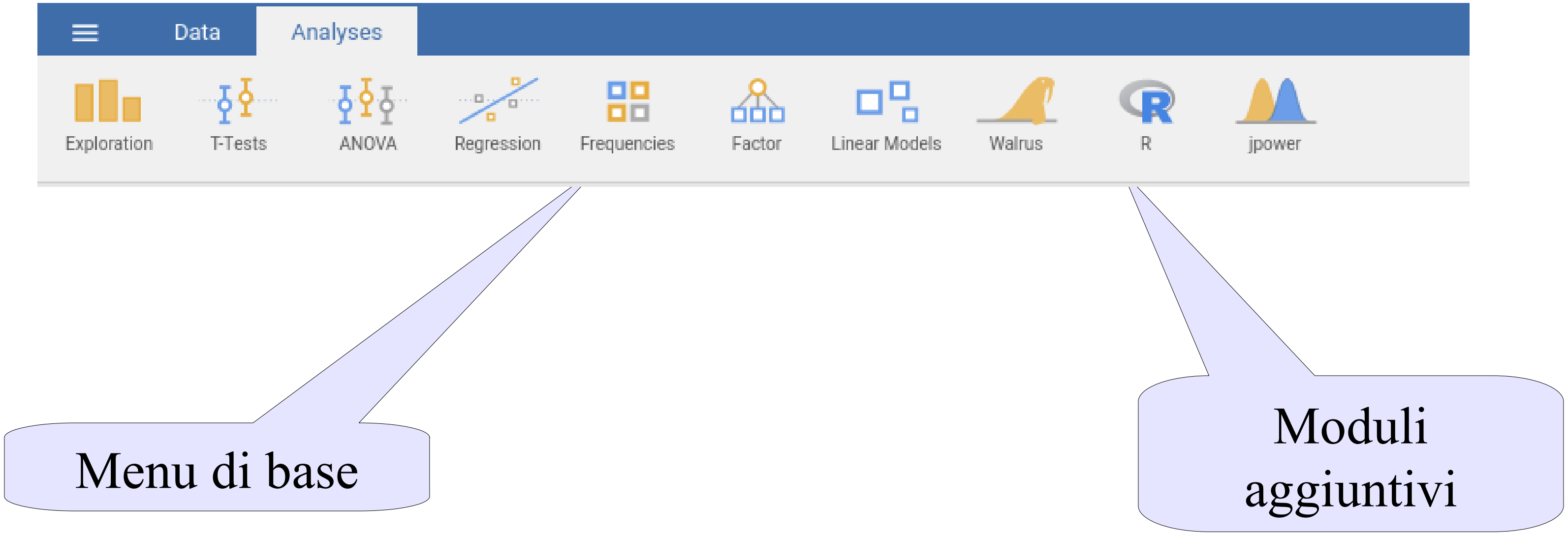
jamovi

- Facciamo queste semplici analisi in **jamovi**



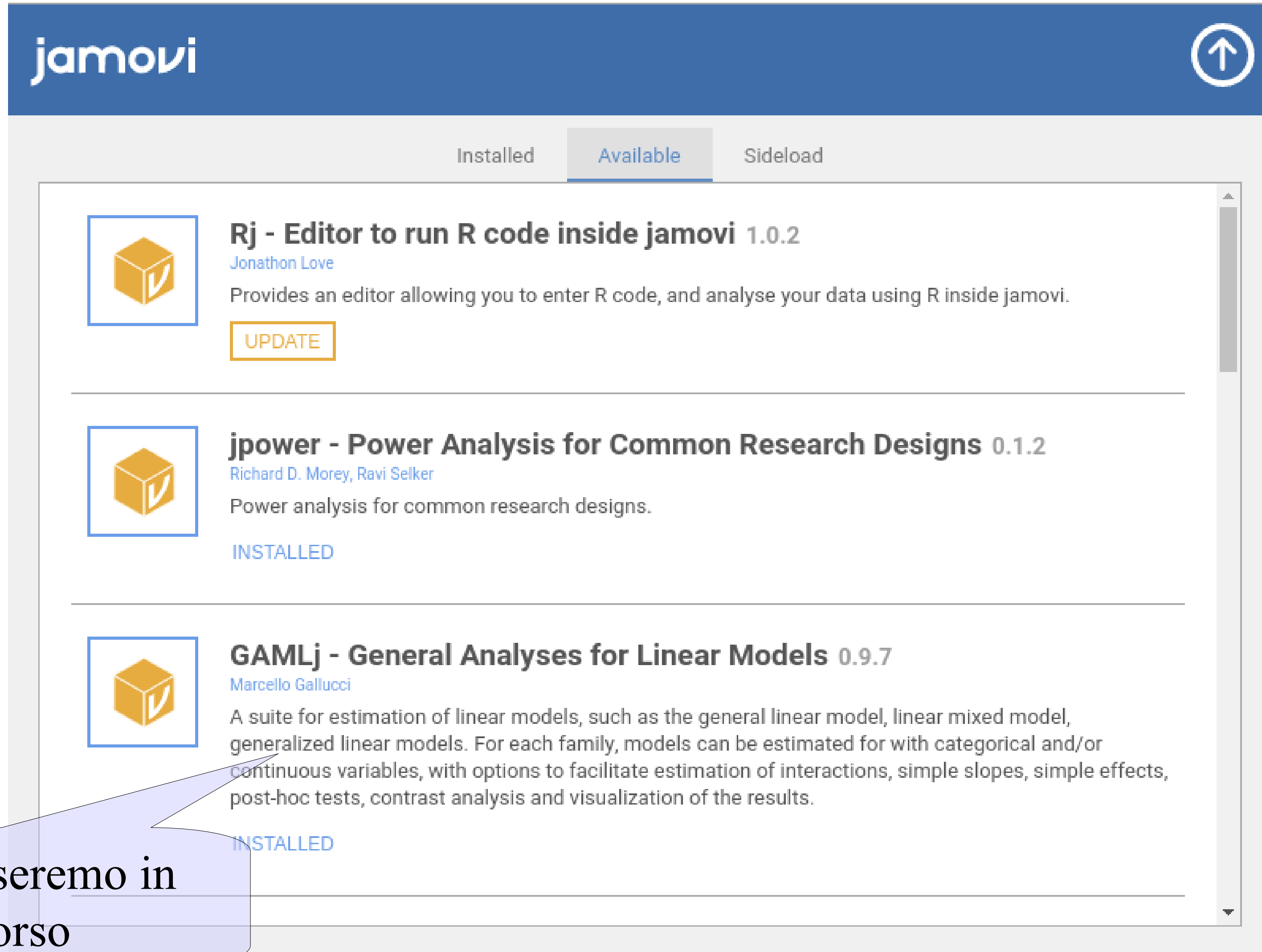
jamovi

- Facciamo queste semplici analisi in **jamovi**



jamovi

- Libreria dei moduli aggiuntivi



Modulo che useremo in questo corso

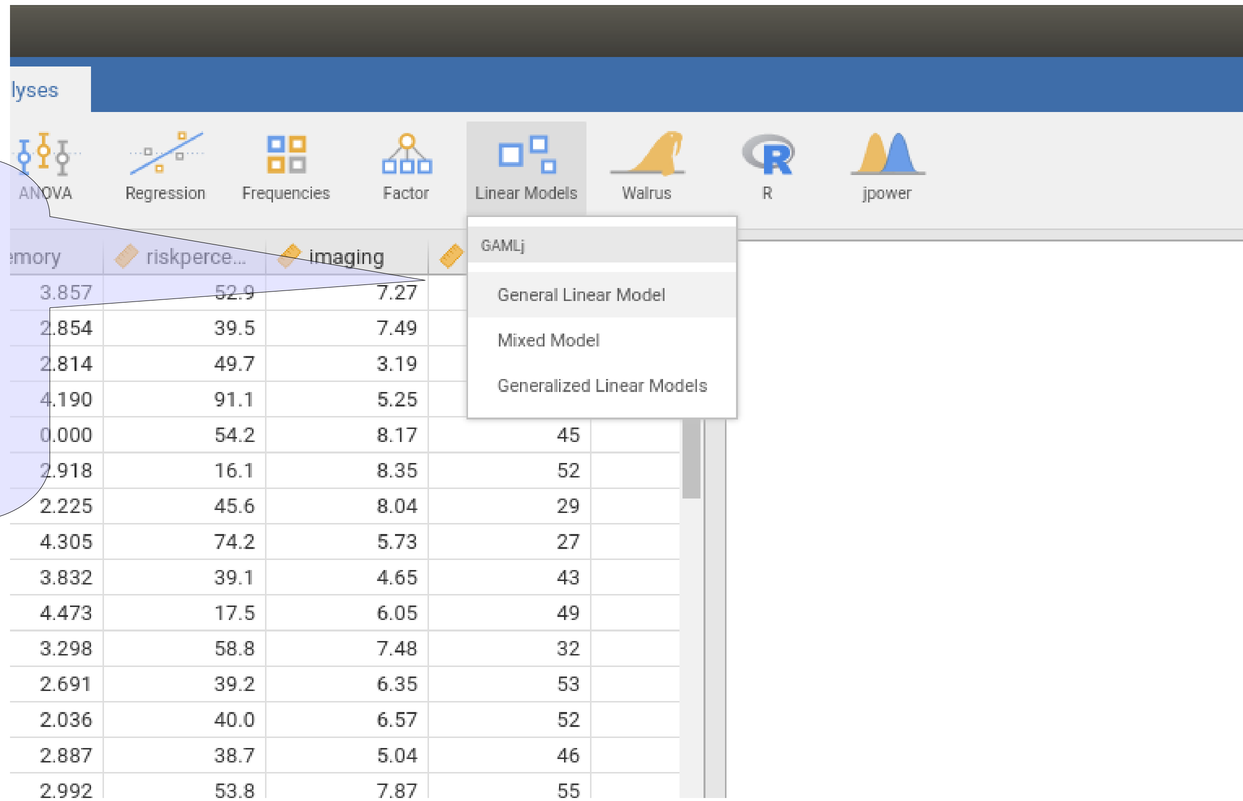
jamovi

- Per la maggior parte dei modelli lineari possiamo usare il modulo GAMLj di jamovi

GLM

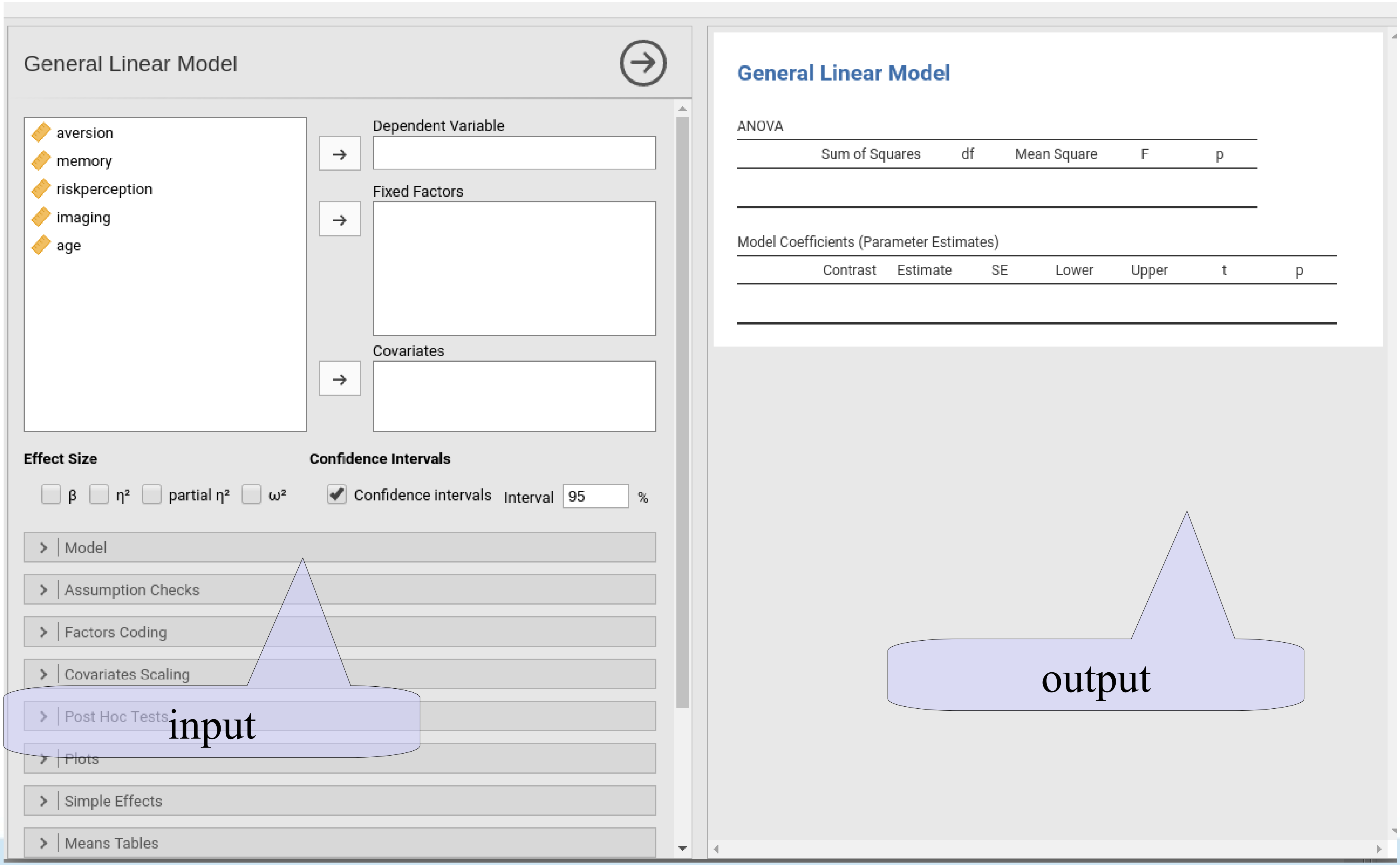
Modelli misti

Modelli gerelizzati



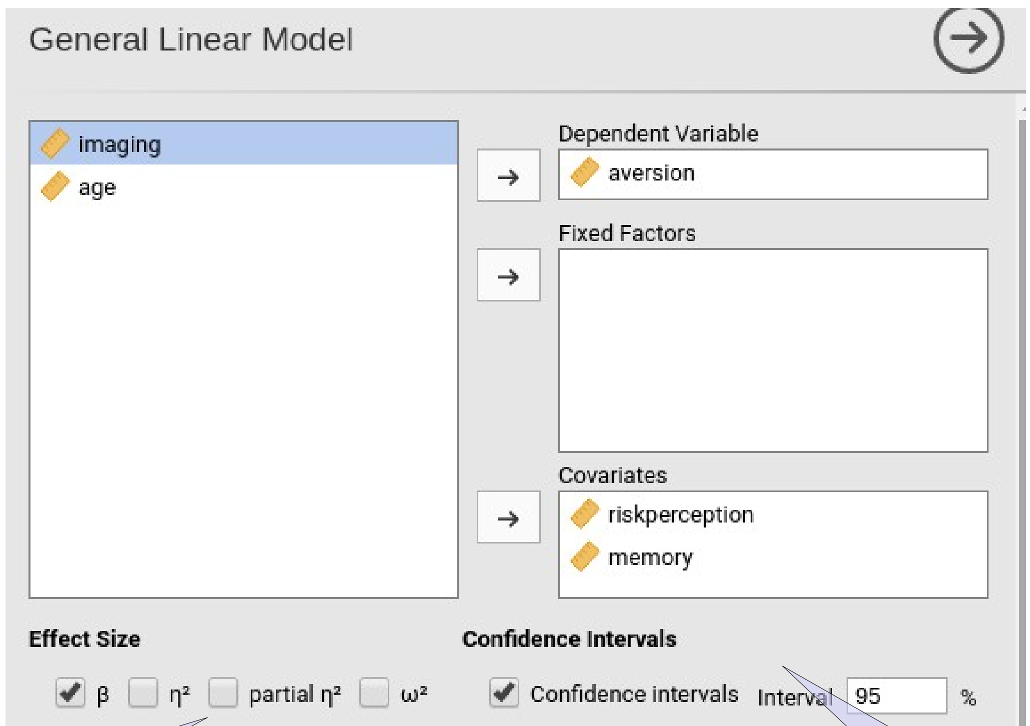
jamovi: interfaccia

- Ogni modulo di jamovi ha la stessa struttura di interfaccia



jamovi: interfaccia

- Ogni modulo di jamovi ha la stessa struttura di interfaccia



Effect sizes

VI continue

jamovi: output

- Come in SPSS, otteniamo i risultati relativi alle varianze (F-test) e ai coefficienti (b, beta, t-test)

Varianze

General Linear Model

ANOVA

	Sum of Squares	df	Mean Square	F	p
Model	86820	2	43410	156.25	< .001
riskperception	78779	1	78779	283.56	< .001
memory	298	1	298	1.07	0.303
Residuals	26949	97	278		

Note. R-squared= 0.763 , adjusted R-squared= 0.758

Model Coefficients (Parameter Estimates)

				95% Confidence Interval				
	Contrast	Estimate	SE	Lower	Upper	t	p	Beta
(Intercept)	Intercept	4.70	1.6668	1.40	8.01	2.82	0.006	0.0000
riskperception	riskperception	1.44	0.0856	1.27	1.61	16.84	< .001	0.8590
memory	memory	1.98	1.9059	−1.81	5.76	1.04	0.303	0.0529

Coefficienti

Variabili Indipendenti Categoriche

ANOVA

Categoriche come IV

- In generale, ed per mediazione e moderazione, è importante capire come il GLM accomoda le variabili indipendenti categoriche

- Consideriamo prima le variabili dicotomiche, cioè con solo due valori (due gruppi)

Esempio

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

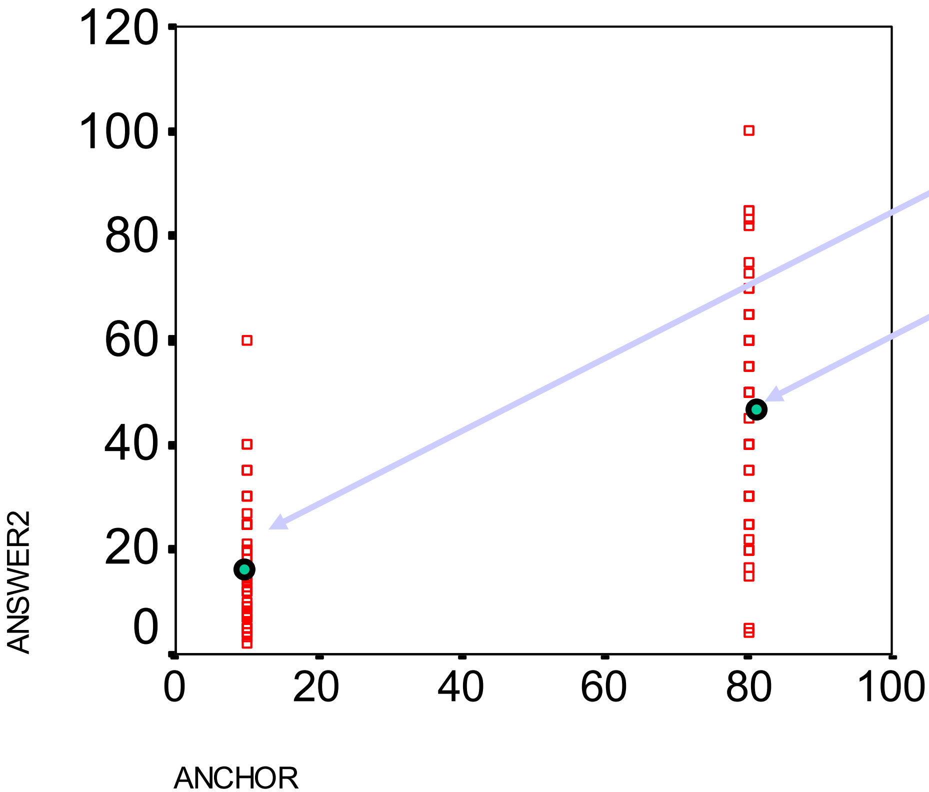
- Domanda 1 a tutti i soggetti: Secondo te, le nazioni africane alle nazioni unite sono più o meno del X %.
- Domanda 2: Quante sono le nazioni africane in percentuale alle nazioni unite
- Gruppo 1: ancora 10%. Gruppo 2: ancora 80%

Scatter Plot

Variabile dipendente: percentuale attesa, variabile indipendente *ancora alta vs bassa*

Medie per gruppo

```
##      0      1
## 17.045 44.829
```



Coefficients for dichotomies

- X= Anchor. Bassa=0 Alta=1

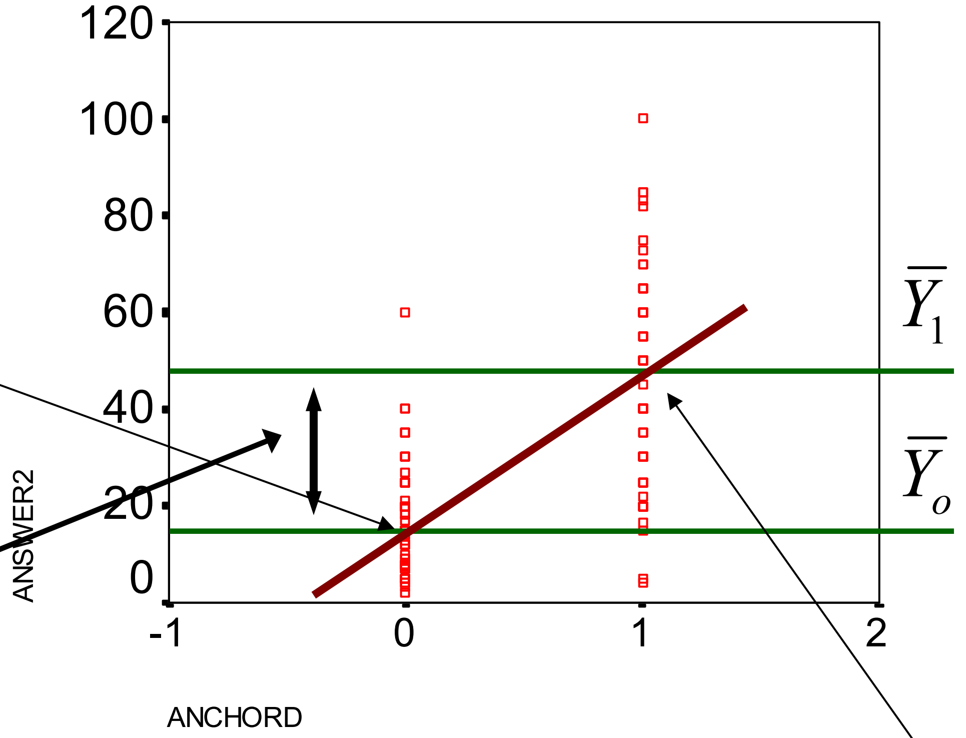
$$y_i = a + b \cdot x_i$$

Termine costante, X=0=media gruppo bassa

$$y_i = a + b \cdot 0 = a = \bar{y}_o$$

b , muovendo X da 0 a 1

$$b = \bar{y}_1 - \bar{y}_o$$



$$y_i = \bar{y}_0 + b \cdot 1 = \bar{y}_1$$

Il B rappresenta la differenza fra le medie dei gruppi

Esempio

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   17.045      1.697   10.04   <2e-16 ***
## groups        27.784      2.400   11.58   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15.64 on 168 degrees of freedom
## Multiple R-squared:  0.4438, Adjusted R-squared:  0.4405
## F-statistic: 134.1 on 1 and 168 DF,  p-value: < 2.2e-16
```

Media gruppo 0

differenza

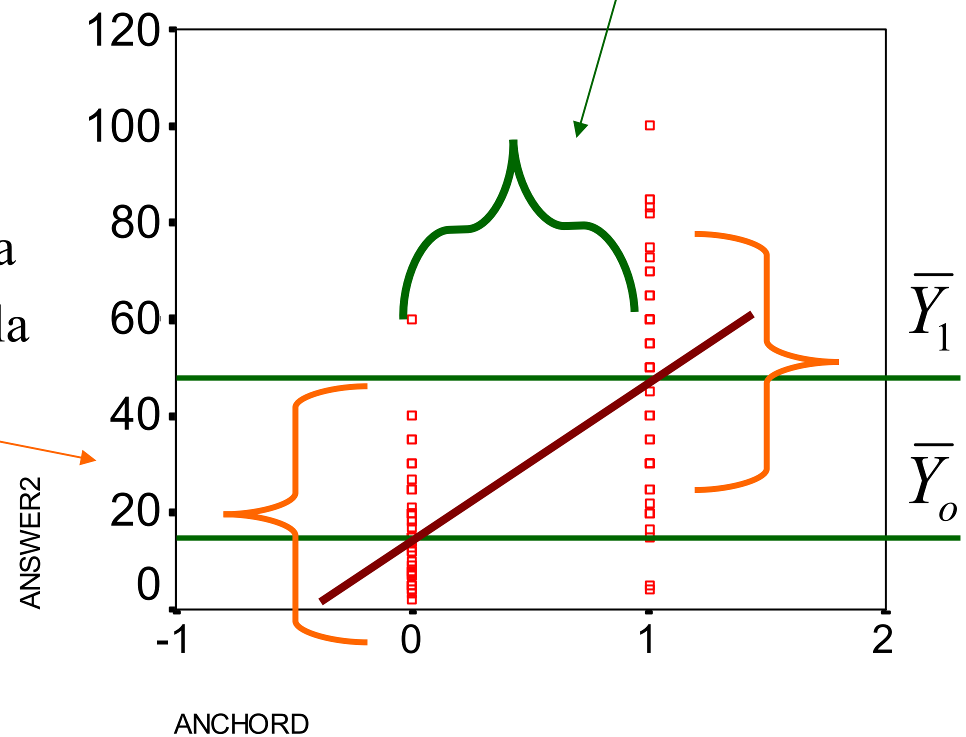
Medie per gruppo

```
##      0      1
## 17.045 44.829
```

R^2 per le dicotomiche

- L' R^2 è la varianza spiegata dalle differenze tra I gruppi (between groups), cioè tra le medie

- La varianza residua 1-R^2 è la varianza non spiegata, cioè la varianza dentro i gruppi (variance within groups)



Test F per dicotomiche

- Il Test F, come per qualunque altro R², p dato da...

$$F = \frac{R^2}{1 - R^2} \frac{df_{within}}{df_{between}} \qquad F = \frac{\text{variance between}}{\text{variance within}} \frac{df_{within}}{df_{between}}$$

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   17.045      1.697   10.04   <2e-16 ***
## groups        27.784      2.400   11.58   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15.64 on 168 degrees of freedom
## Multiple R-squared:  0.4438, Adjusted R-squared:  0.4405
## F-statistic: 134.1 on 1 and 168 DF,  p-value: < 2.2e-16
```

Test F per dicotomiche

- Possiamo anche chiedere direttamente il test F per l'effetto (utile quando ci sono più effetti), ottendendo cosi l'ANOVA

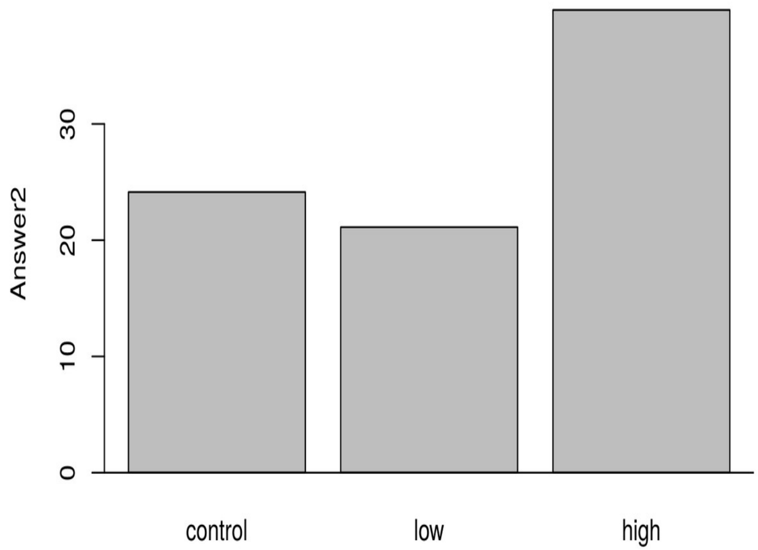
```
## Analysis of Variance Table
##
## Response: answer2
##            Df Sum Sq Mean Sq F value    Pr(>F)
## groups      1  32808   32808  134.06 < 2.2e-16 ***
## Residuals 168  41113     245
## ---
```

Più di due categorie

- Quando si hanno più di due categorie, si rappresentano le variabili mediante una serie di dummy variables
- Una dummy è una variabile dicotomica
- Consideriamo un esempio come il precendente, ma con tre gruppi: Ancora bassa, Ancora alta, e no Ancora

Medie per gruppo

```
##      0     1     2
## 24.14 21.12 39.80
```



Più di due categorie

- L'informazione contenuta in una variabile nominale (K>2) può essere rappresentata da un numero K-1 variabili dicotomiche

- K-1 variabili dicotomiche è il numero minore di dicotomiche in grado di rappresentare i gruppi

Queste variabili sono dette dummies

Possiamo distinguere i gruppi? Gruppi: Control, Low, High

Variabile	Categoria	var1	var2
	Control	0	0
Groups	Low	1	0
	High	0	1

3 gruppi, 2 dummies
K gruppi, K-1 dummies

Coefficienti per le dummies

- Se usiamo queste variabili in una regressione…

$$Y = a + B_1 \cdot \begin{array}{c} \text{var1} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \end{array} + B_2 \cdot \begin{array}{c} \text{var2} \\ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{array} \quad \begin{array}{r} \text{Control} \\ \text{Low} \\ \text{High} \end{array}$$

Cosa è il termine costante **a**?

Il valore medio atteso di DV per tutte le dummies uguali a zero

$$Y = a + B_1 \cdot 0 + B_2 \cdot 0 = a = \bar{Y}_{control}$$

Coefficienti per le dummies

- Cosa è il B associato a var1?

$$Y = a + B_1 \cdot \begin{array}{c} \text{var1} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \end{array} + B_2 \cdot \begin{array}{c} \text{var2} \\ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{array} \quad \begin{array}{r} \text{Control} \\ \text{Low} \\ \text{High} \end{array}$$

Cosa è il coefficiente B1?

$$Y = \bar{Y}_{control} + B_1 \cdot Low + B_2 \cdot 0$$

$$B_1 = \bar{Y}_{Low} - \bar{Y}_{Control}$$

Differenza tra Low e Control

Coefficienti per le dummies

- Cosa è il B associato a var2?

$$Y = a + B_1 \cdot \begin{array}{c} \text{var1} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \end{array} + B_2 \cdot \begin{array}{c} \text{var2} \\ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{array} \quad \begin{array}{r} \text{Control} \\ \text{Low} \\ \text{High} \end{array}$$

Cosa è il coefficiente B2?

$$Y = \bar{Y}_{control} + B_1 \cdot 0 + B_2 \cdot High$$

$$B_2 = \bar{Y}_{High} - \bar{Y}_{Control}$$

Differenza tra High e Control

Dummies

- Chiameremo il gruppo che ha tutti zero nelle dummies (Control) il reference group
- La constante della regressione è la media della DV per il reference group
- Il B di ogni dummy representa la differenza tra il gruppo con dummy=1 e il reference group
- Il test di significatività di ogni B testa che tale differenza sia diversa da zero

Esempio

- R codifica le variabili categoriche mediante il comando `factor()`.

- Nel nostro esempio le dummy risultano

Medie per gruppo

```
##     0     1     2
## 24.14 21.12 39.80
```

Dummies

```
##         groups1 groups2
## control       0       0
## low           1       0
## high          0       1
```

Esempio

- Chiamando il comando per il modello lineare otteniamo i risultati

```
## Call:
## lm(formula = answer2 ~ groups, data = anchor2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.800   -7.800    0.030    6.625   30.200
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   24.140      1.539  15.683  < 2e-16 ***
## groups1       -3.020      2.177  -1.387    0.167
## groups2       15.660      2.177   7.194 2.97e-11 ***
## ---
```

Low vs Control

High vs Control

Dummies

```
##          groups1 groups2
## control        0       0
## low            1       0
## high           0       1
```

Esempio: Coefficienti (parametri)

- Chiamando il comando per il modello lineare otteniamo i risultati

```
## Call:
## lm(formula = answer2 ~ groups, data = anchor2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.800   -7.800    0.030    6.625   30.200
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   24.140      1.539  15.683  < 2e-16 ***
## groups1       -3.020      2.177  -1.387    0.167
## groups2       15.660      2.177   7.194 2.97e-11 ***
## ---
```

Low vs Control

High vs Control

Medie per gruppo

```
## control     low    high
##   24.14   21.12   39.80
```

Esempio: F-test

- Volendo, si può ottenere anche la F-test aggregata, cioè dell'effetto principale

```
## Analysis of Variance Table
##
## Response: answer2
##            Df Sum Sq Mean Sq F value    Pr(>F)
## groups      2  10055  5027.5  42.441 2.823e-15 ***
## Residuals 147  17413   118.5
## ---
```

Differenti codifiche

- E' possibile codificare le variabili dummies in molti modi diversi
- A seconda di come vengono codificate le dummies, l'interpretazione dei coefficienti cambia

Contrast (deviation) coding

Possiamo distinguere i gruppi? Gruppi: Control, Low, High

Variabile	Categoria	var1	var2
	Control	-1	-1
Groups	Low	1	0
	High	0	1

Le dummies hanno tutte media 0: sono centrate sulla media

Coefficienti per contrast (deviation) coding

$$Y = a + B_1 \cdot \overset{\text{var1}}{\begin{bmatrix}-1\\ 1\\ 0\end{bmatrix}} + B_2 \cdot \overset{\text{var2}}{\begin{bmatrix}-1\\ 0\\ 1\end{bmatrix}} \quad \begin{array}{r}\text{Control}\\ \text{Low}\\ \text{High}\end{array}$$

- Cosa è l'intercetta? La media di Y nel campione totale
- Cosa è B1? La differenza tra la media nel campione e la media del gruppo Low
- Cosa è B2? La differenza tra la media nel campione e la media del gruppo High

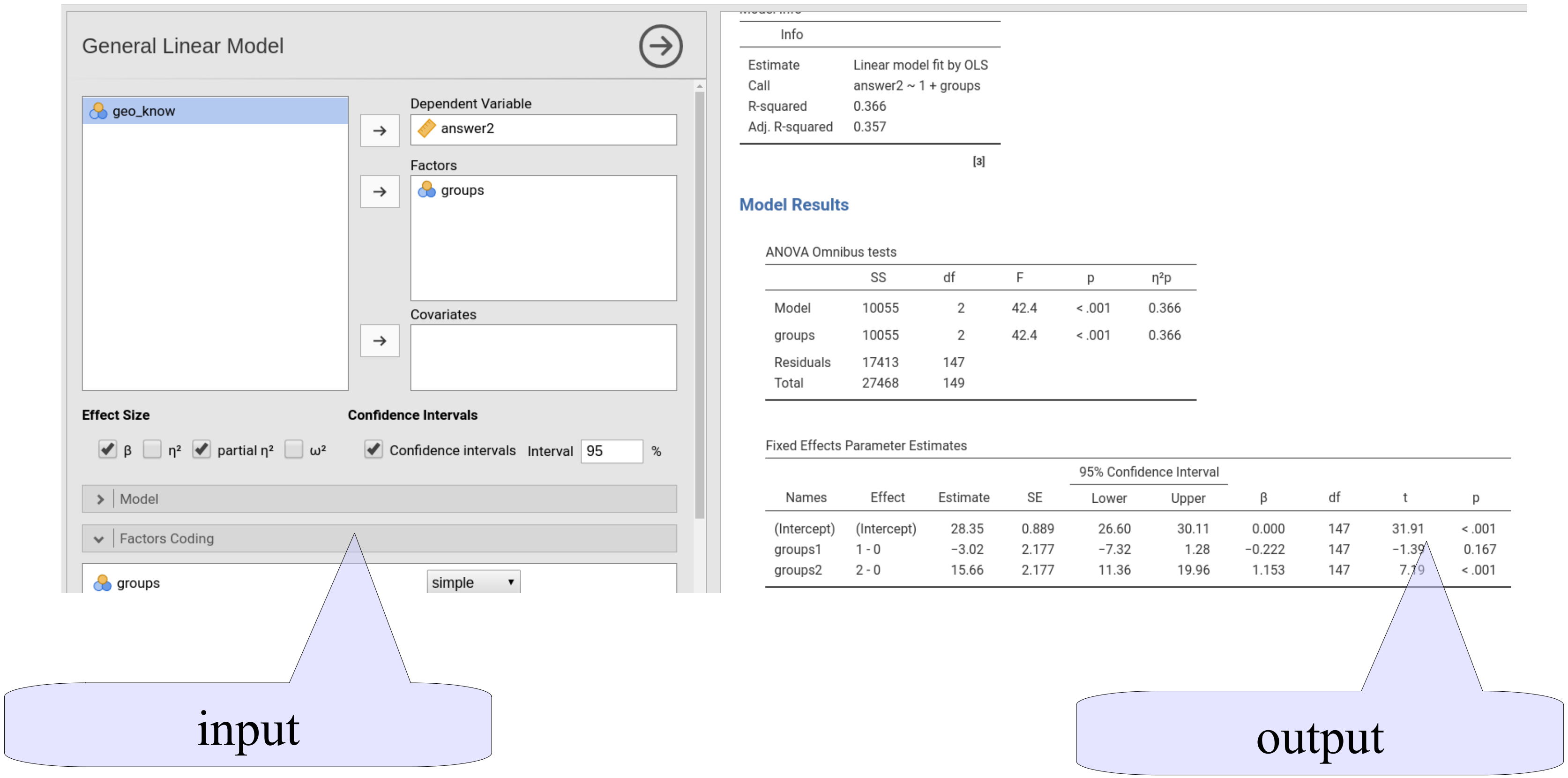
Coefficienti per contrast (simple) coding

$$Y = a + B_1 \cdot \begin{array}{c} \text{var1} \\ \begin{bmatrix} -.333 \\ .666 \\ -.333 \end{bmatrix} \end{array} + B_2 \cdot \begin{array}{cc} \text{var2} & \\ \begin{bmatrix} -.333 \\ -.333 \\ .666 \end{bmatrix} & \begin{array}{l} \text{Control} \\ \text{Low} \\ \text{High} \end{array} \end{array}$$

- Cosa è l'intercetta? La media di Y nel campione totale
- Cosa è B1? Il confronto tra gruppo Control e gruppo Low
- Cosa è B2? Il confronto tra gruppo Control gruppo e gruppo High
- **"Simple" è come "dummy" ma centrato sullo zero**

jamovi: ANOVA

- GAMLj di jamovi semplifica il tutto



ANOVA Omnibus tests

	SS	df	F	p	η²p
Model	10055	2	42.4	< .001	0.366
groups	10055	2	42.4	< .001	0.366
Residuals	17413	147			
Total	27468	149			

Fixed Effects Parameter Estimates

				95% Confidence Interval					
Names	Effect	Estimate	SE	Lower	Upper	β	df	t	p
(Intercept)	(Intercept)	28.35	0.889	26.60	30.11	0.000	147	31.91	< .001
groups1	1 - 0	-3.02	2.177	-7.32	1.28	-0.222	147	-1.39	0.167
groups2	2 - 0	15.66	2.177	11.36	19.96	1.153	147	7.19	< .001

Jamovi: ANOVA

- GAMLj di jamovi semplifica il tutto

Model Results

Effetto principale

ANOVA Omnibus tests

	SS	df	F	p	η²p
Model	10055	2	42.4	< .001	0.366
groups	10055	2	42.4	< .001	0.366
Residuals	17413	147			
Total	27468	149			

Fixed Effects Parameter Estimates

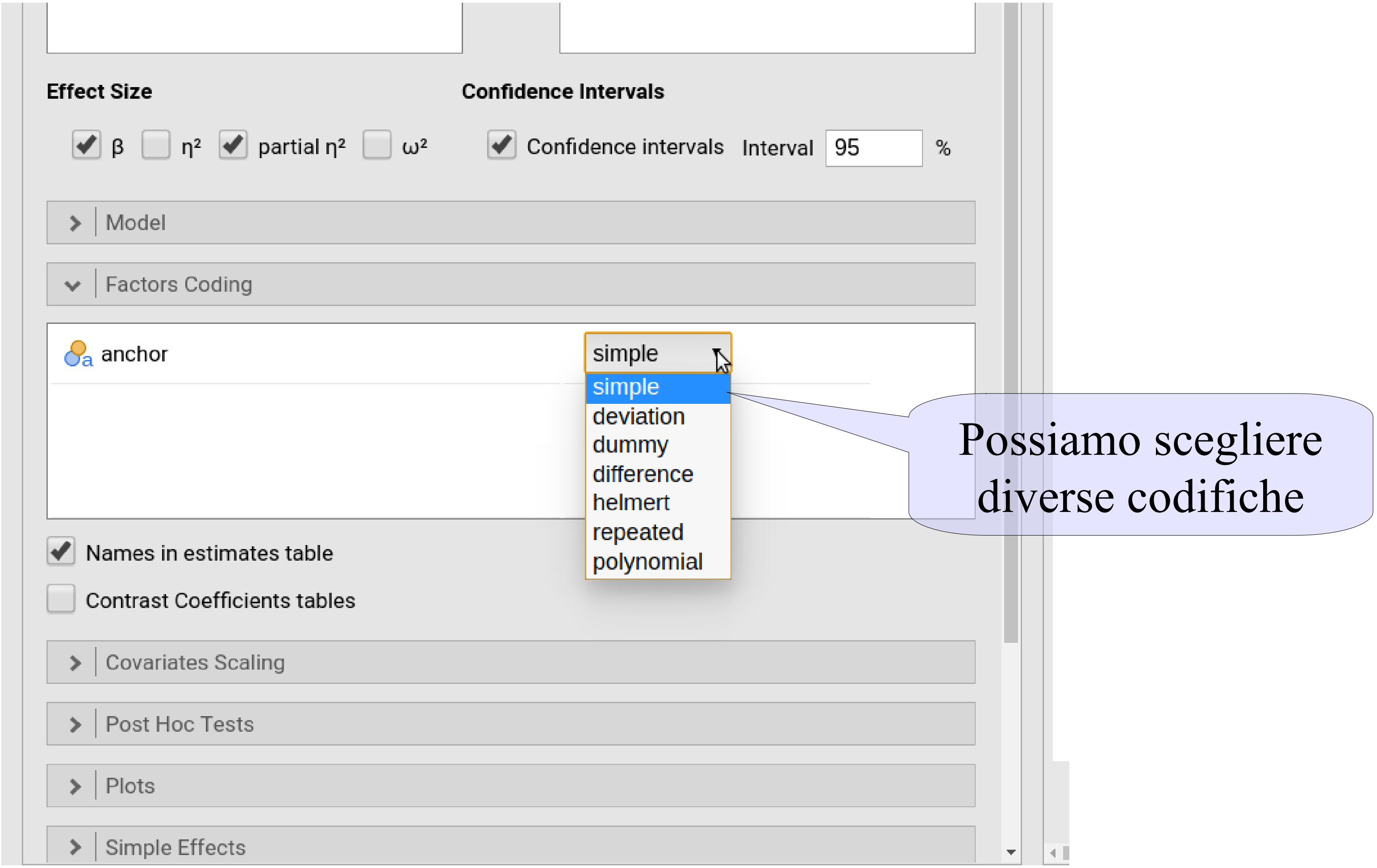
				95% Confidence Interval					
Names	Effect	Estimate	SE	Lower	Upper	β	df	t	p
(Intercept)	(Intercept)	28.35	0.889	26.60	30.11	0.000	147	31.91	< .001
groups1	1 - 0	-3.02	2.177	-7.32	1.28	-0.222	147	-1.39	0.167
groups2	2 - 0	15.66	2.177	11.36	19.96	1.153	147	7.19	< .001

Etichette dei confronti

parametri

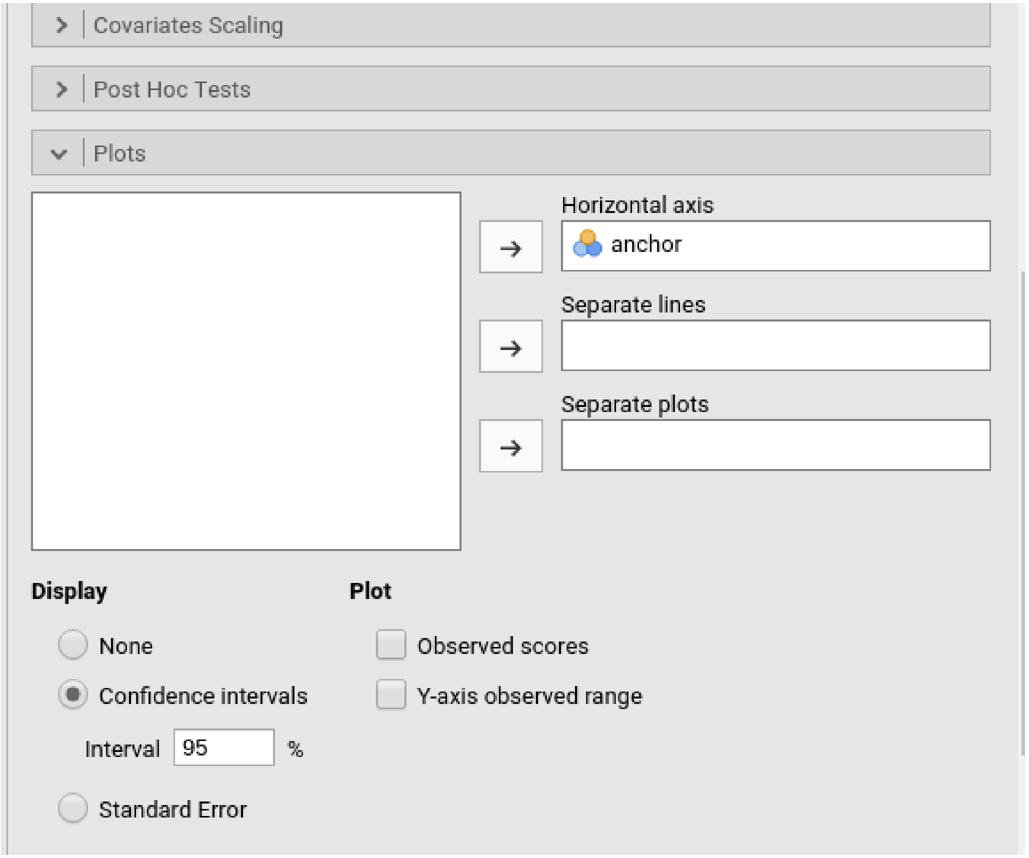
Jamovi: ANOVA

- GAMLj: differenti codifiche per i contrast



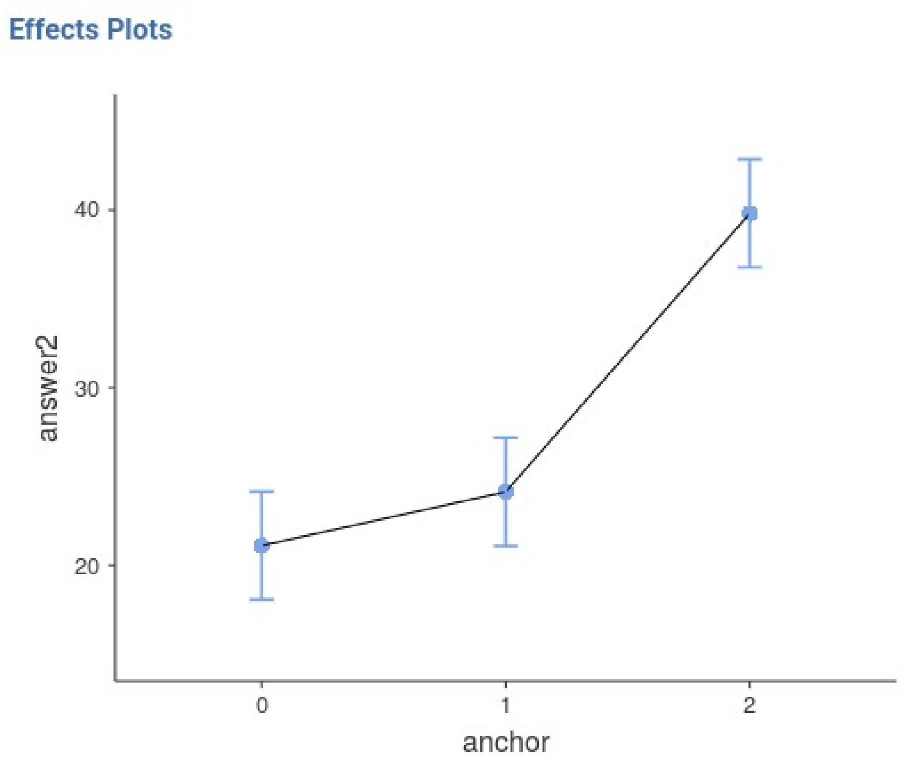
Jamovi: ANOVA

- GAMLj: grafici delle medie



Jamovi: ANOVA

- GAMLj: grafici delle medie (e intervalli di confidenza)



GLM

- A questo punto:
- Ogni cosa che si impara a fare con le variabili indipendenti continue può essere applicato alle categoriche (mediante dummies)
- Ogni cosa che si può fare con le VI categoriche si può fare con le VI continue

Effect size indices

GLM

- La grandezza dell'effetto si può valutare in termini di
  - Dei coefficienti
  - Varianza spiegata (associata ad un effetto)

Effect size: Coefficienti

- Consideriamo un modello con *aversion* come VD

Effect size non standardizzato

Effect size standardizzato

Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval Lower	95% Confidence Interval Upper	β	df	t	p
(Intercept)	(Intercept)	4.725	1.676	1.398	8.051	0.000	96	2.819	0.006
memory	memory	1.879	1.941	−1.974	5.732	0.050	96	0.968	0.335
age - Cat1	1 - 0	1.029	3.398	−5.716	7.774	0.030	96	0.303	0.763
riskperception	riskperception	1.443	0.086	1.272	1.614	0.860	96	16.752	< .001

Effect size: Coefficienti standardizzati

- Consideriamo un modello con *aversion* come VD

VI continue: di quante deviazioni standard varia Y al variare di X al netto delle altre VI

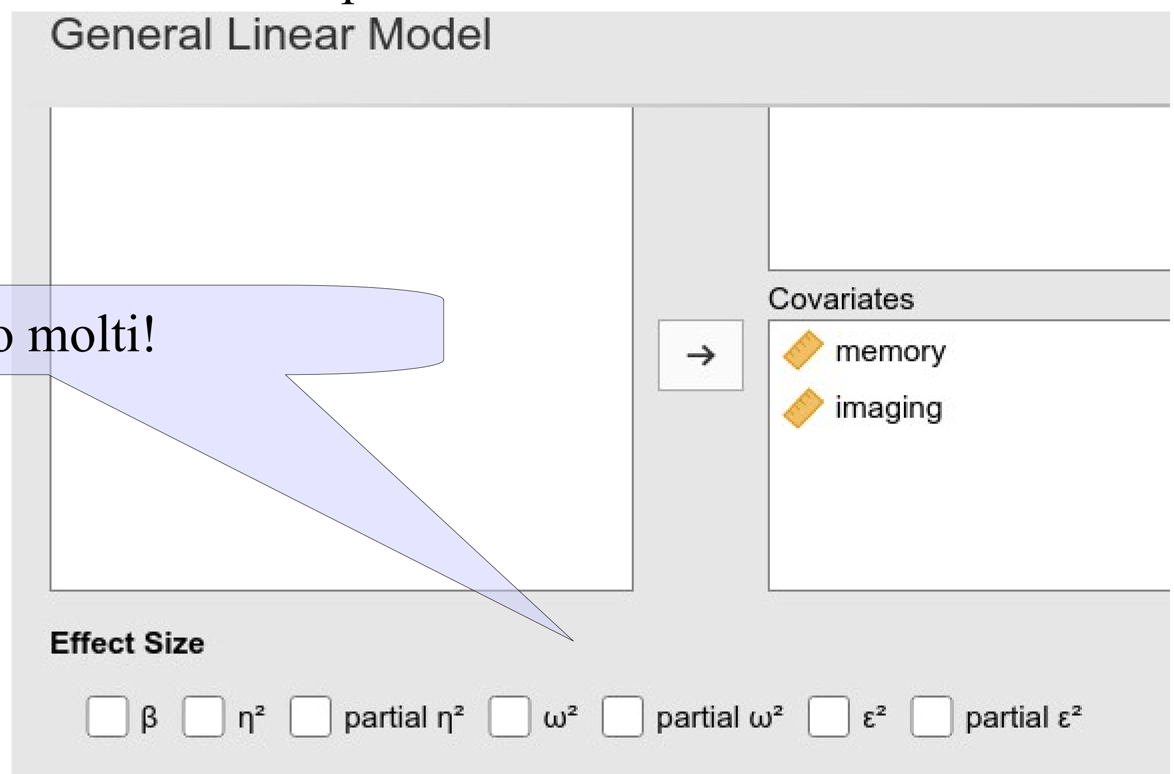
Fixed Effects Parameter Estimates

				95% Confidence Interval					
Names	Effect	Estimate	SE	Lower	Upper	β	df	t	p
(Intercept)	(Intercept)	4.725	1.676	1.398	8.051	0.000	96	2.819	0.006
memory	memory	1.879	1.941	−1.974	5.732	0.050	96	0.968	0.335
age - Cat1	1 - 0	1.029	3.398	−5.716	7.774	0.030	96	0.303	0.763
riskperception	riskperception	1.443	0.086	1.272	1.614	0.860	96	16.752	< .001

VI categoriche: Di quante deviazioni standard di Y i due gruppi differiscono, al netto delle altre VI

Effect size: Varianze

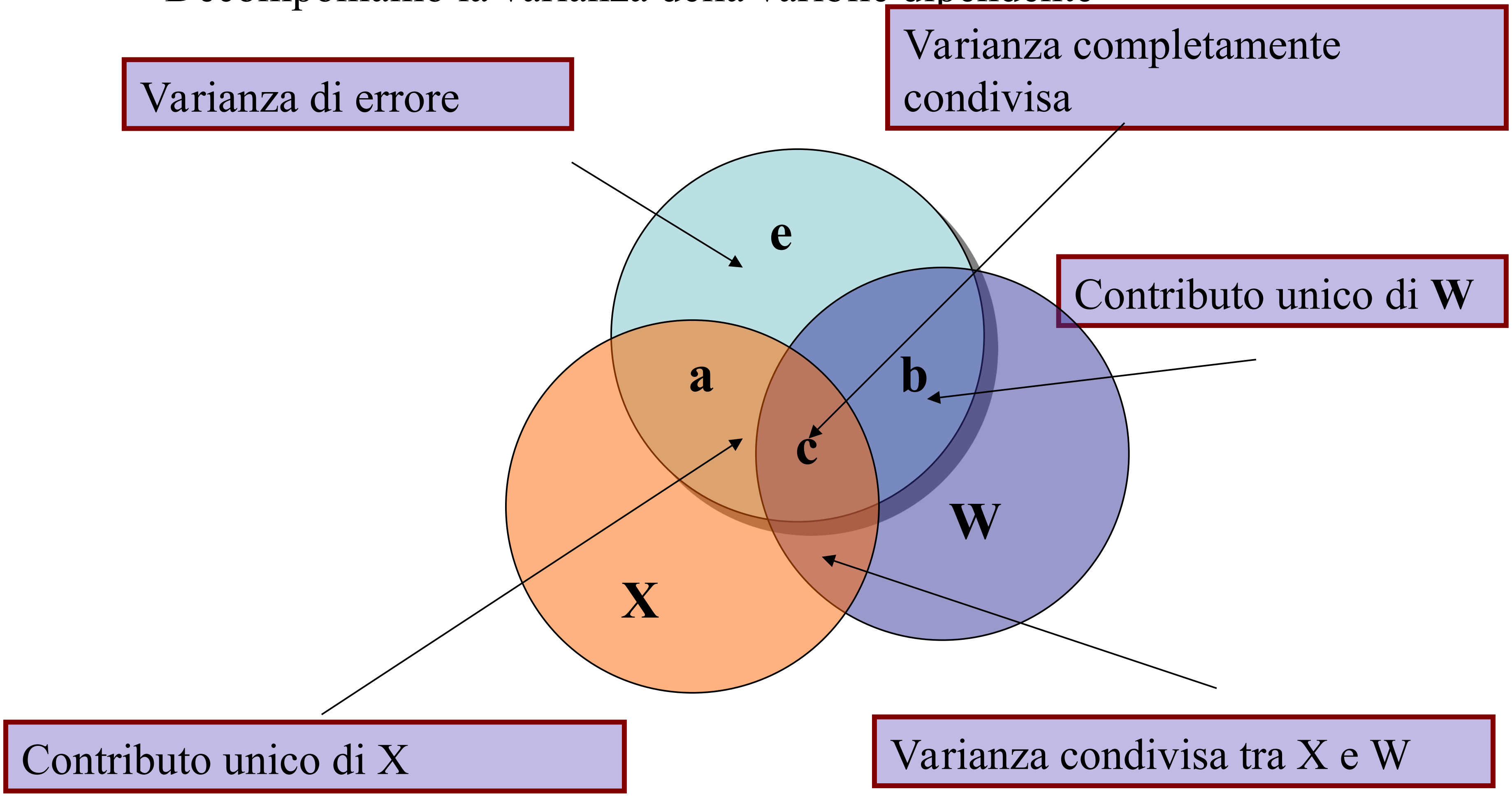
- Gli effect size indices basati sulla varianza spiegata indicano quanto ogni effetto (ogni variabile indipendente) contribuisce a spiegare varianza della variabile dipendente



Ce ne sono molti!

Varianza Decomposta

- Decomponiamo la varianza della varibile dipendente

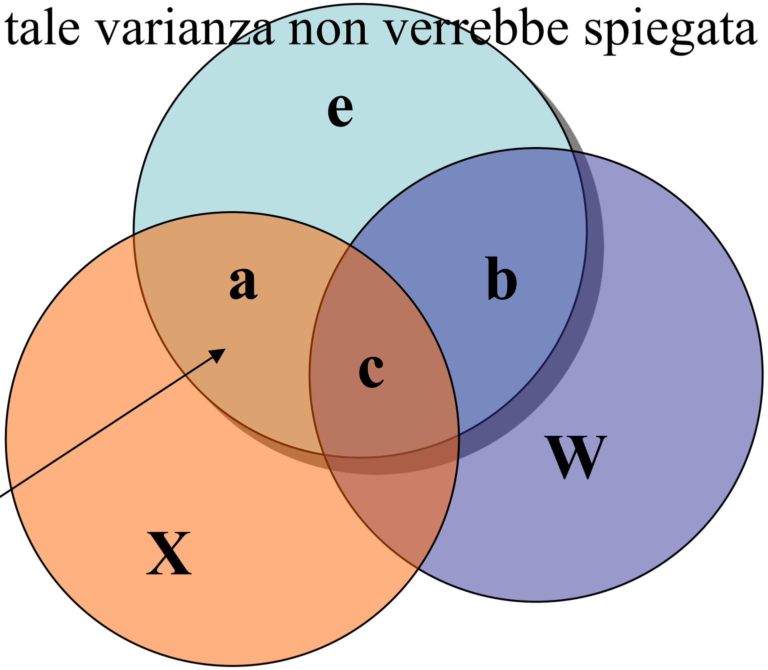


Contributo unico

- Il contributo unico di ogni variabile è la varianza unicamente spiegata dalla variabile.

- In assenza di quella variabile, tale varianza non verrebbe spiegata

$$CU = a$$



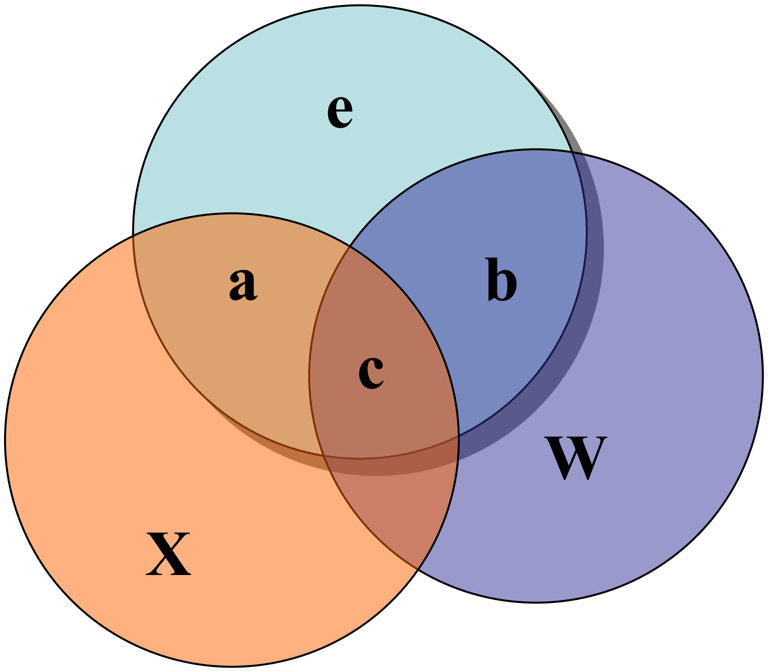
Contributo unico di X

Proporzione di varianza

- CU deve essere espresso come una proporzione di varianza

Eta-squared

$$\eta^2 = \frac{a}{a+b+c+e}$$



Partial eta-squared

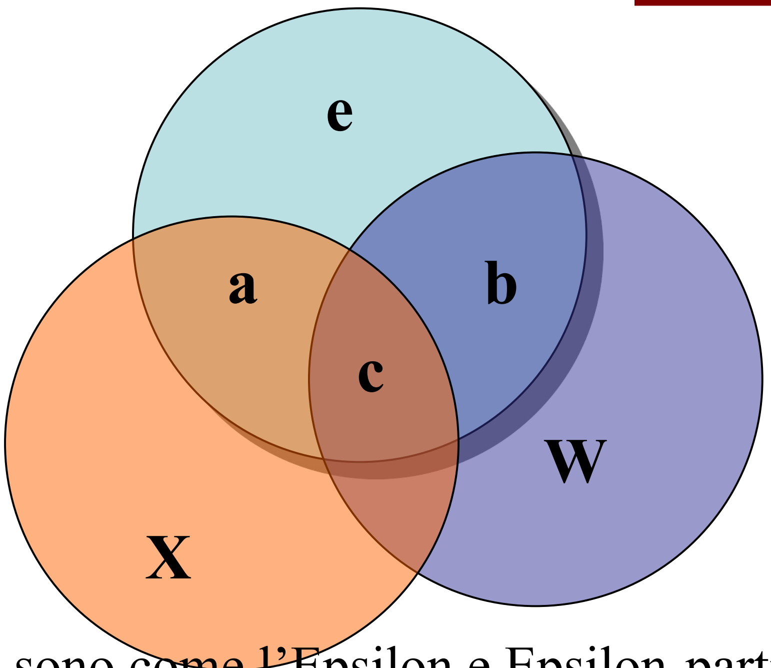
$$\eta^2_\partial = \frac{a}{a+e}$$

Proporzione di varianza

- Gli Eta sono calcolati sul campione, ma le varianze possono essere stimate nella popolazione (adjusted)

Epsilon-squared

$$\epsilon^2 = E\left(\eta^2\right)$$



Partial Epsilon-squared

$$\epsilon^2 = E\left(\eta_\partial^2\right)$$

- Omega e Omega-partial sono come l'Epsilon e Epsilon-partial, la formula di stima è leggermente diversa ma equivalente in pratica

Proporzione di varianza

- Gli Eta sono calcolati sul campione, ma le varianze possono essere stimate nella popolazione (adjusted)

	Campione	Popolazione
Full	Eta-squared	Epsilon-squared
Partial	Partial-Eta squared	Partial Epsilon-squared

Effect size: Coefficienti

- Consideriamo un modello con *aversion* come VD

ANOVA Omnibus tests

	SS	df	F	p	η^2	$\eta^2 p$	ε^2	$\varepsilon^2 p$
Model	86845.412	3	103.222	< .001	0.763	0.763	0.756	0.756
memory	262.902	1	0.937	0.335	0.002	0.010	−0.000	−0.001
age - Cat	25.719	1	0.092	0.763	0.000	0.001	−0.002	−0.009
riskperception	78698.096	1	280.614	< .001	0.692	0.745	0.689	0.742
Residuals	26923.146	96						
Total	113768.558	99						

Effect size: Morale

- Eta (full) è più onesto dell'Eta-squared, anche se la maggior parte degli autori riporta l'Eta-squared
- Epsilon (e Omega) sono stime più accurate delle varianze, dunque andrebbero preferiti
- Riportate l'Epsilon, ma aspettatevi che i reviewers vi chiedano l' partial Eta-squared

GLM: Morale

- Il modello lineare generale consente di stimare gli effetti tra variabili dipendenti continue e variabili indipendenti categorico o continue
- Sulla base di questi coefficienti è possibile modellare la stima del modello di mediazione e di moderazione, o di qualunque altra combinazione di modelli
- Ciò nelle prossime lezioni