

Analisi Multivariata dei Dati

Regressione Multipla (cap. 3)

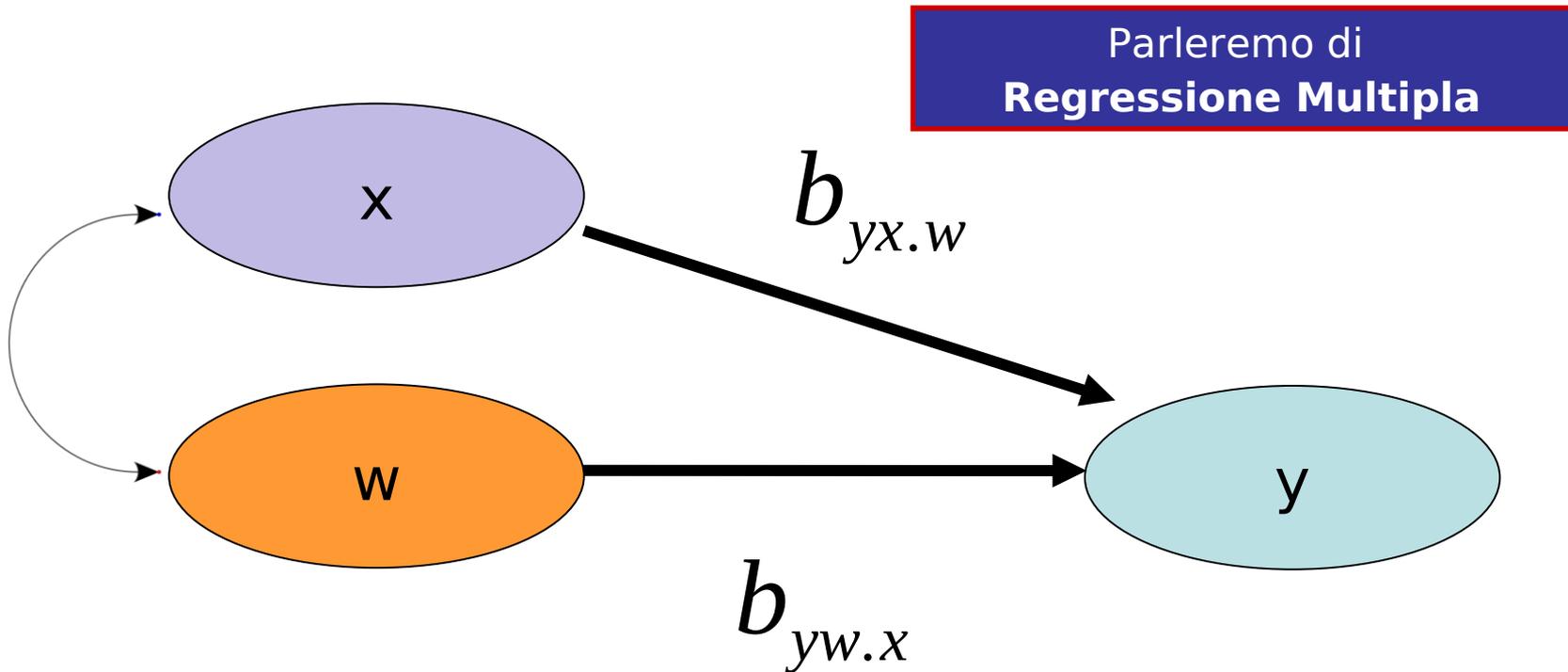
Marcello Gallucci

Milano-Bicocca

A
M
D

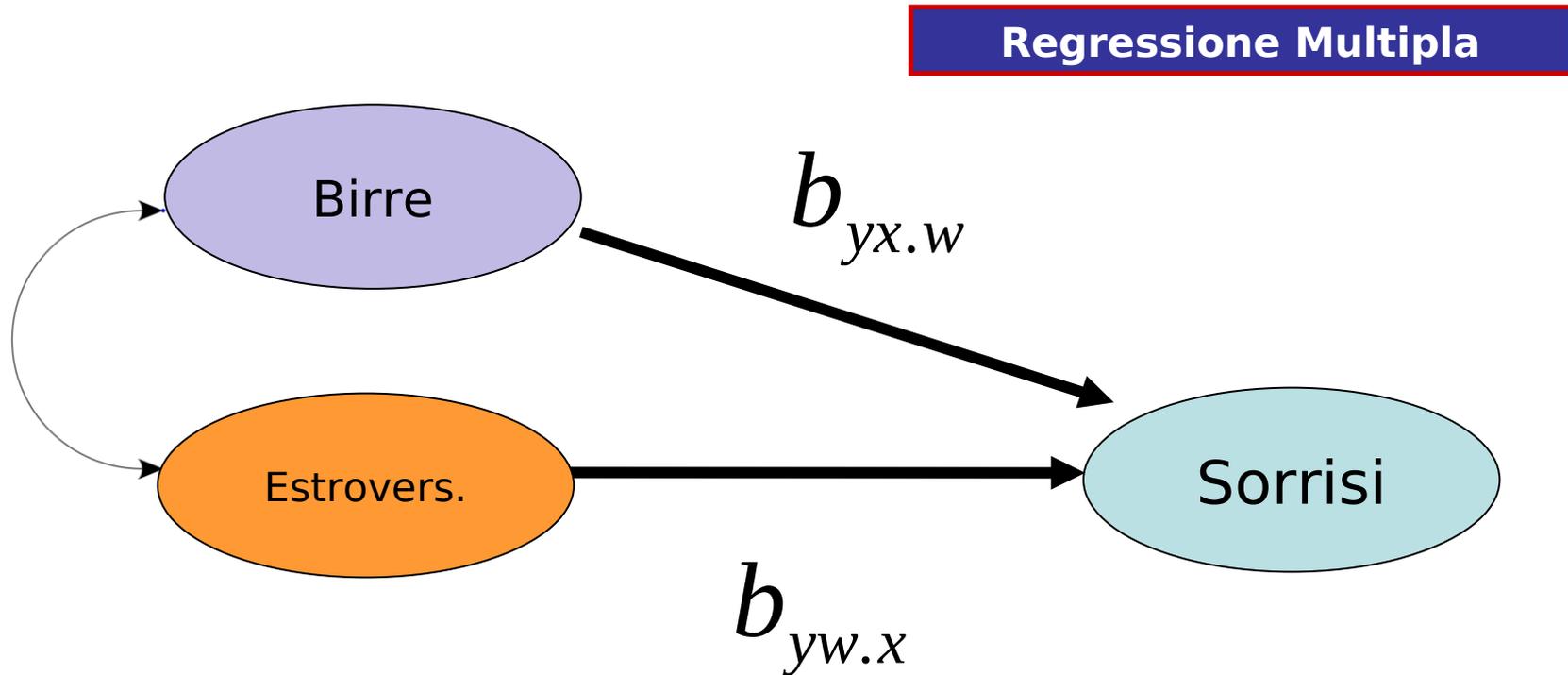
Effetti multipli

- ◆ Il caso in cui la variabile dipendente possa essere spiegata da più di una variabile



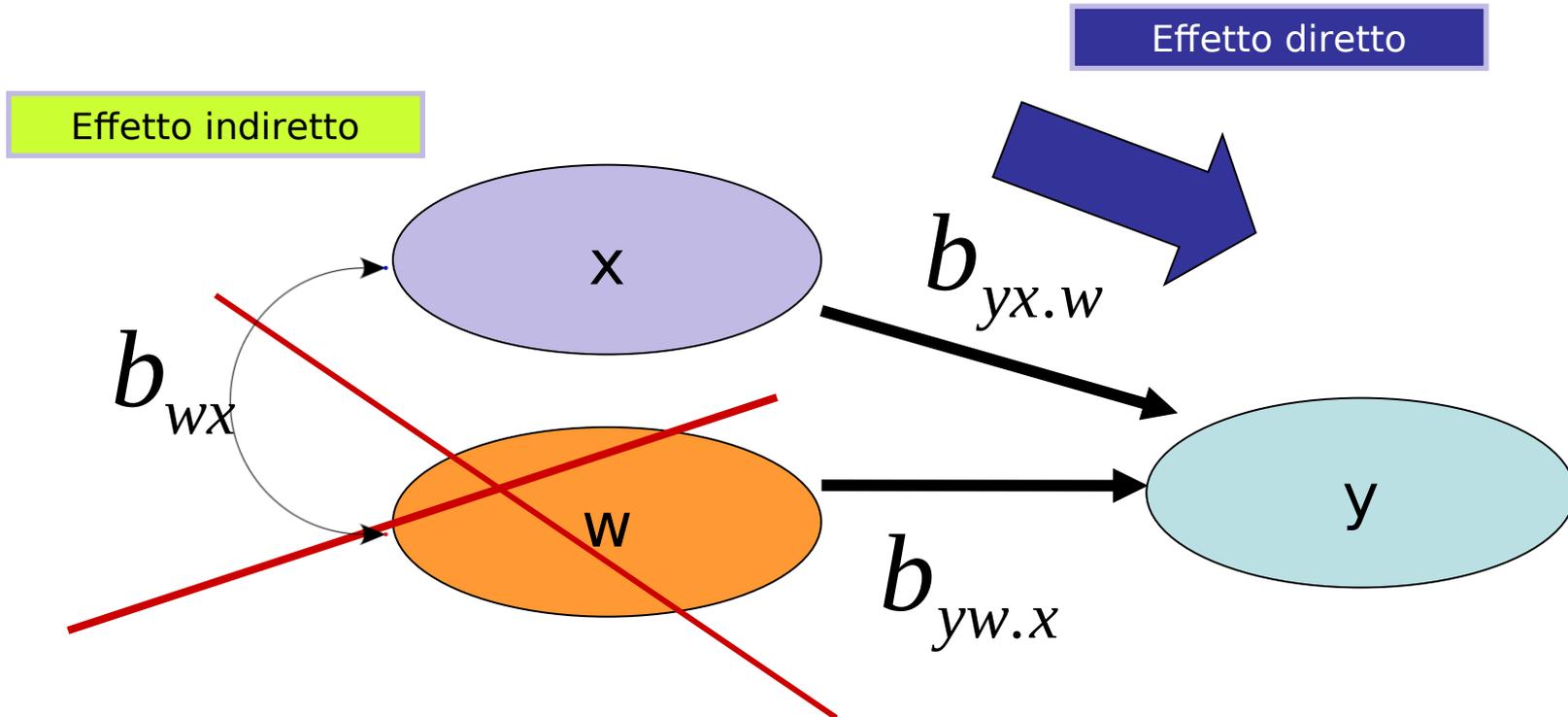
Esempio Effetti multipli

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



Effetti Parziali

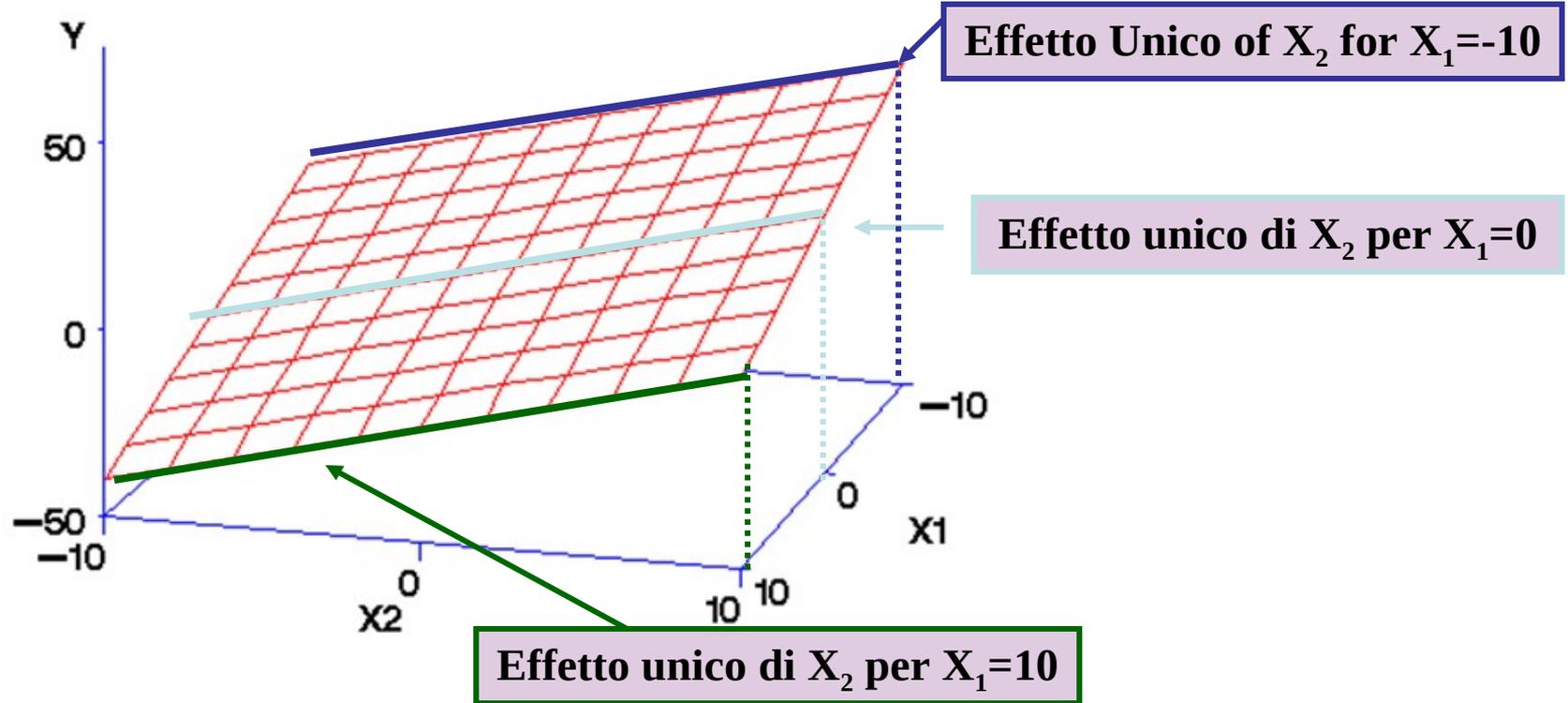
- Togliere l'effetto indiretto è equivalente a bloccare la possibilità che x vada su y mediante w : Il coefficiente viene dunque detto **coefficiente parziale**, cioè l'effetto di x parzializzando l'effetto di w



Interpretazione geometrica

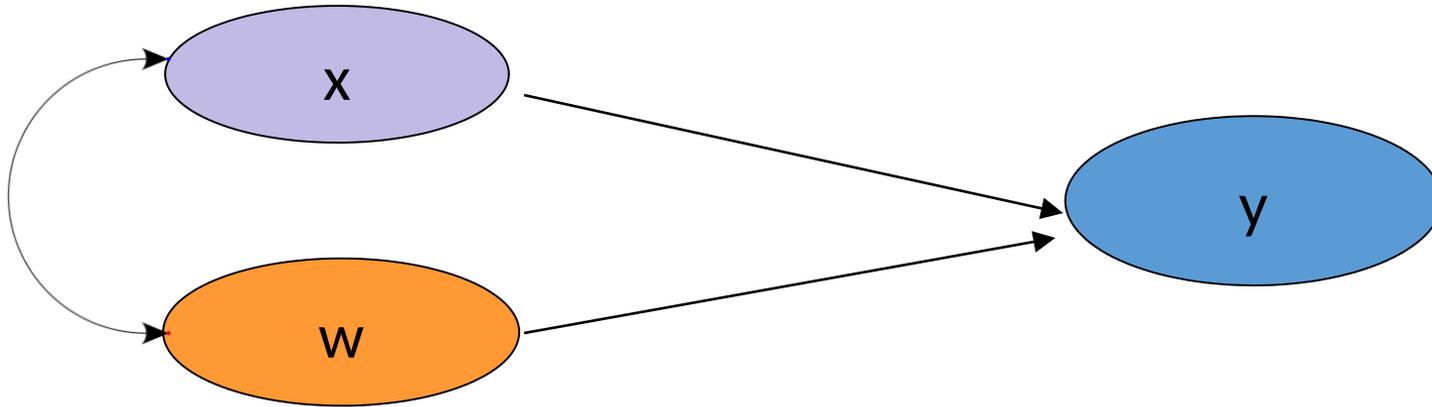
Regression surface

$$Y = -2 \cdot X_1 + 2 \cdot X_2$$



Perché la regressione multipla?

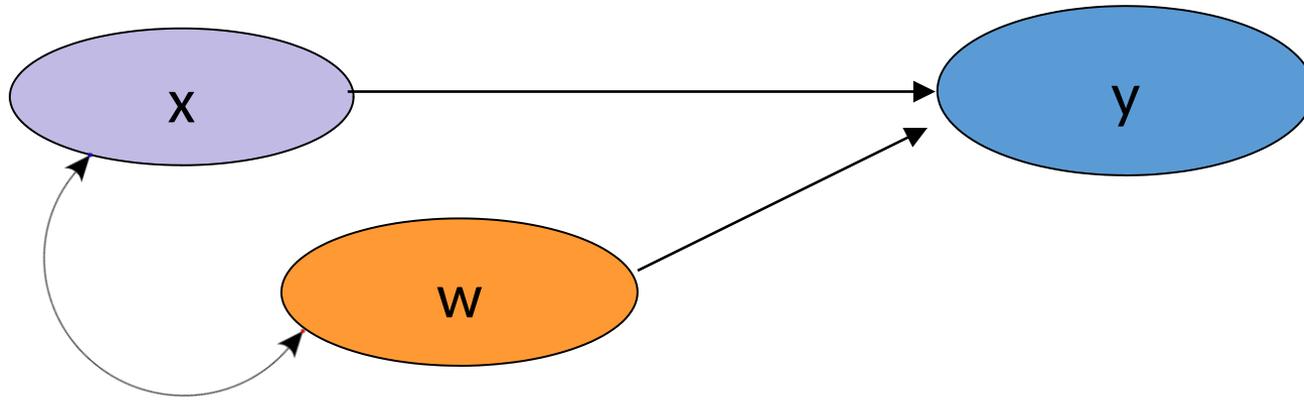
- Per testare modelli con più di un predittore



- ◆ Es. Effetto delle competenze linguistiche e dell'autocontrollo sull'abilità di lettura dei bambini/e
- ◆ Ma molto spesso i predittori hanno varianza in comune (sono correlati)
- ◆ La regressione multipla **isola il legame unico** di ogni predittore con il criterio

Perché la regressione multipla?

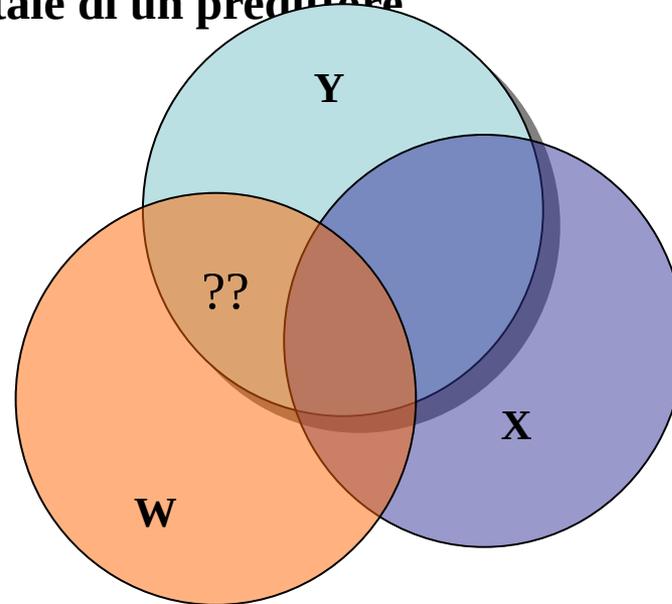
- ◆ Per tenere sotto controllo la variabilità nel criterio causata da variabili che non sono per noi teoricamente interessanti, ma possono creare ‘rumore’ nei dati



- ◆ Effetto della pubblicità sull'intenzione di acquistare un prodotto, tenendo sotto controllo altri fattori importanti (es. età del consumatore, genere, reddito)

Perché la regressione multipla?

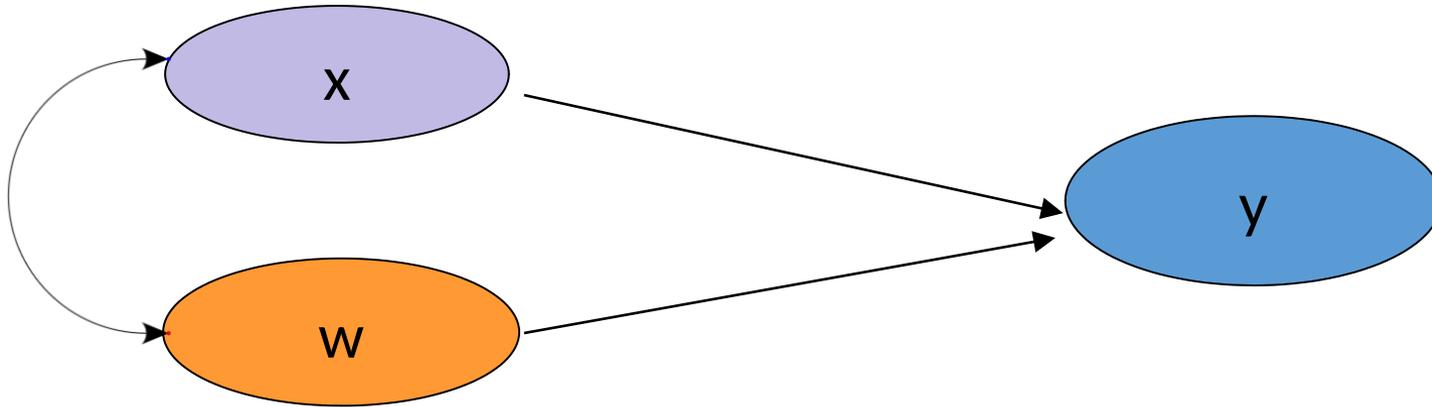
- ◆ Per stimare la **validità incrementale di un predittore**



- ◆ Una nuova variabile è predittiva dopo aver tenuto sotto controllo altre variabili?
- ◆ Quanto spiega **la percezione dei ricchi del fumo** dell'intenzione di smettere in più rispetto al prezzo delle sigarette e le condizioni di salute del soggetto

Perché la regressione multipla?

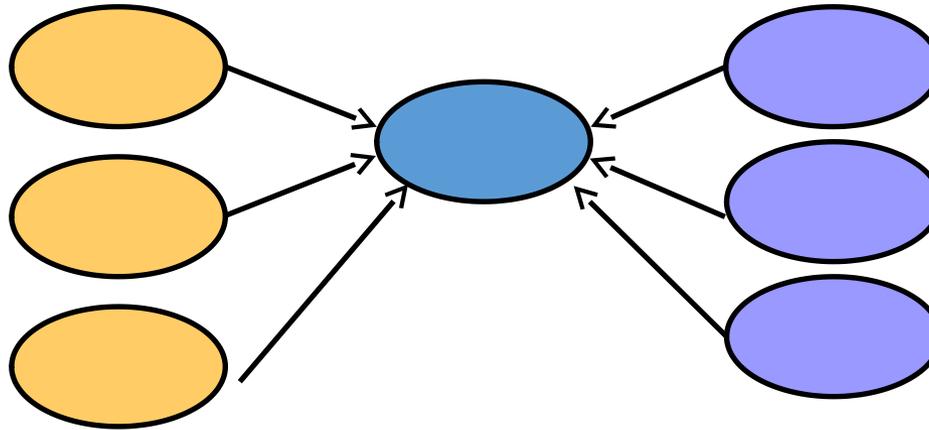
- Per testare quali predittori sono più importanti?



- ◆ Es. quale predittore del voto di laurea è maggiormente predittivo: il voto di maturità, il test di ingresso,
- ◆ Selezionare i predittori migliori e più utili

Perché la regressione multipla?

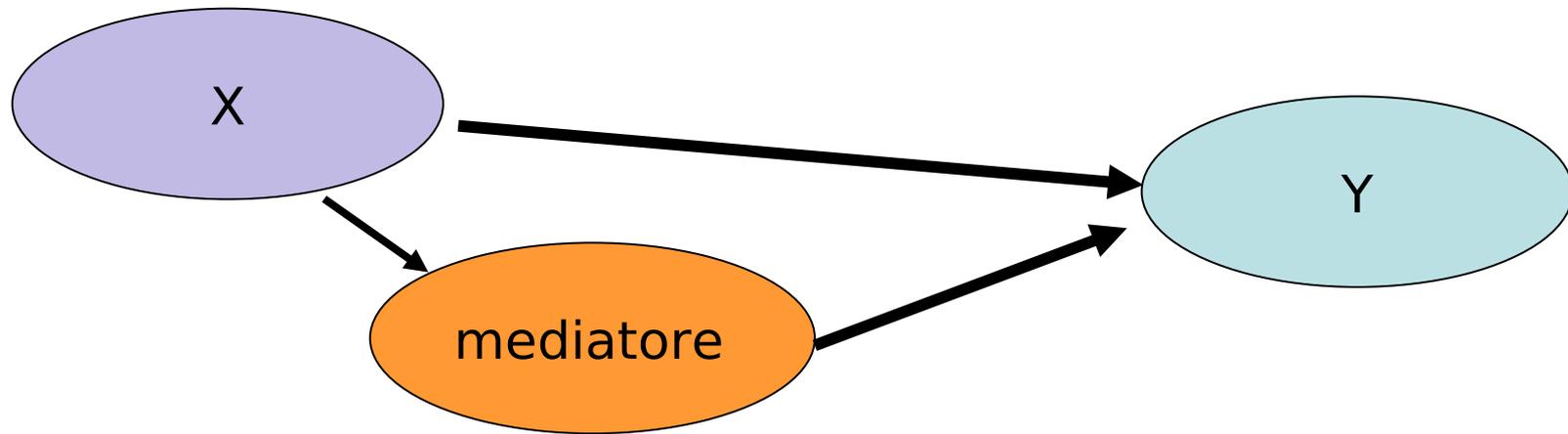
- Per confrontare gruppi di variabili indipendenti



- ◆ Es. Cosa è più importante nel determinare il grado di felicità di una persona, gli aspetti sociali (es. numero di amici e contatti) oppure gli aspetti materiali (es. reddito, proprietà, status)?

Perché la regressione multipla?

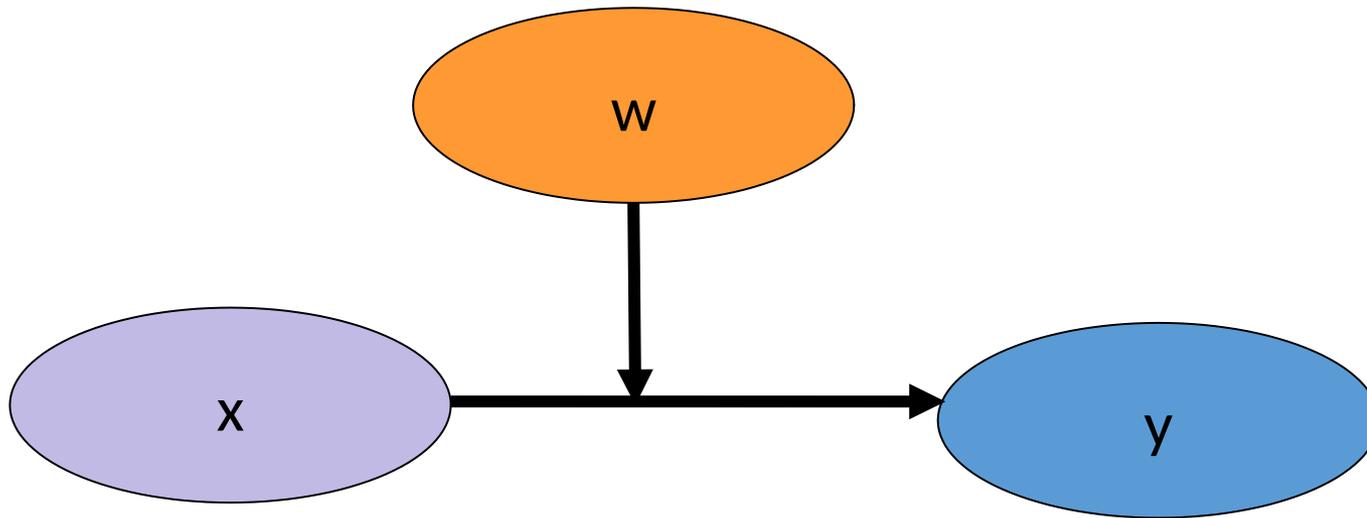
- ◆ Per testare modelli teorici più complessi
- ◆ Per esempio modelli di **mediazione**



- ◆ Le persone che leggono più romanzi hanno risultati scolastici migliori perché hanno una maggiore padronanza linguistica

Perché la regressione multipla?

- ◆ Per testare modelli teorici più complessi
- ◆ Per esempio modelli di **moderazione**



- ◆ L'effetto della lettura sulle competenze linguistiche è più forte nei giovani che negli anziani

Esempio

◆ Abbiamo rilevato su una classe di 30 studenti universitari il voto preso all'esame di statistica e due variabili che vogliamo usare come predittori del voto: il numero di *ore di studio* per preparare l'esame e il numero di esami sostenuti nell'anno (*appelli*).

Statistiche descrittive

	N	Minimo	Massimo	Media	Deviazione std.
voto	30	12	30	22.17	4.942
appelli	30	0	11	4.67	2.631
ore	30	0	90	38.87	20.714
Numero di casi validi (listwise)	30				

Esempio

- ◆ Incominciamo con una regressione semplice tra *voto* e *ore di studio*

Coefficienti di regressione

Coefficienti^a

Modello		Coefficienti non standardizzati		Coefficienti standardizzati		Sign.
		B	Errore std.	Beta	t	
1	(Costante)	16.861	1.622		10.393	.000
	ore	.137	.037	.572	3.692	.001

a. Variabile dipendente: voto

Esempio

- ◆ Incominciamo con una regressione semplice tra *voto* e *ore di studio*

Riepilogo del modello

Modello	R	R-quadrato	R-quadrato adattato	Errore std. della stima
1	.572 ^a	.327	.303	4.124

a. Predittori: (costante), ore

ANOVA^a

Modello		Somma dei quadrati	gl	Media quadratica	F	Sign.
1	Regressione	231.886	1	231.886	13.632	.001 ^b
	Residuo	476.280	28	17.010		
	Totale	708.167	29			

a. Variabile dipendente: voto

b. Predittori: (costante), ore

Esempio

- ◆ Ora passiamo ad una regressione multiple

Coefficienti

Modello		Coefficienti non standardizzati		Coefficienti standardizzati	t	Sign.
		B	Errore std.	Beta		
1	(Costante)	18.542	2.923		6.343	.000
	appelli	-.236	.339	-.125	-.694	.494
	ore	.122	.043	.510	2.820	.009

a. Variabile dipendente: voto

Esempio

- Incominciamo con una regressione semplice tra *voto* e *ore di studio*

Varianza spiegata

Riepilogo del modello

Modello	R	R-quadrato	R-quadrato adattato	Errore std. della stima
1	.582 ^a	.339	.290	4.163

a. Predittori: (costante), ore, appelli

ANOVA^a

Modello		Somma dei quadrati	gl	Media quadratica	F	Sign.
1	Regressione	240.235	2	120.118	6.931	.004 ^b
	Residuo	467.931	27	17.331		
	Totale	708.167	29			

a. Variabile dipendente: voto

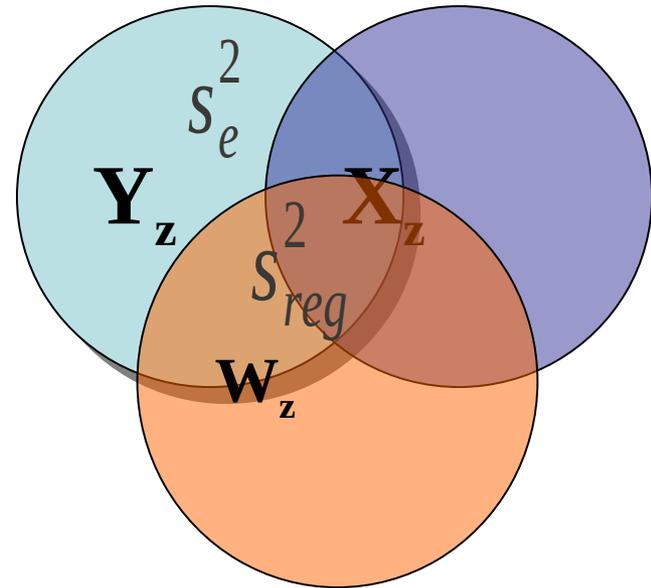
b. Predittori: (costante), ore, appelli

Esempio

- ◆ Concluderemo che le ore di studio hanno un effetto sul voto finale, anche al netto del numero di appelli fatti nel semestre
- ◆ Il numero di appelli, al netto delle ore studiate, non hanno effetto

Varianza spiegata

- ◆ La sostanza e' la stessa che nella regressione semplice



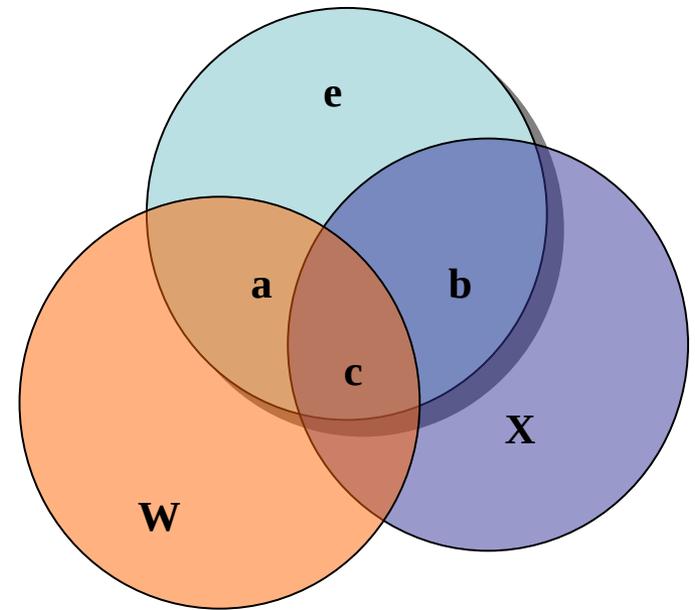
$$\frac{S_{reg}^2}{S_y^2} = \frac{S_y^2 - S_e^2}{S_y^2} = R_{y.xw}^2$$

Varianza non spiegata

- ◆ Percentuale di varianza di errore, non spiegabile mediante la regressione

$$e = 1 - R^2_{y.xw}$$

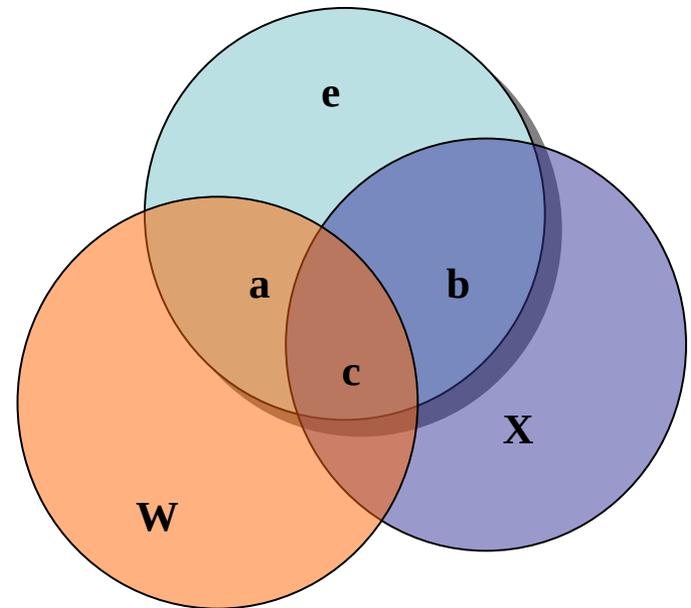
Coefficiente di alienazione



Calcolo di R^2

- ◆ Il calcolo della varianza spiegata puo' essere effettuato partendo dalle correlazioni semplici

$$R_{y.xw}^2 = \frac{r_{yx}^2 + r_{yw}^2 - 2r_{yx}r_{yw}r_{wx}}{1 - r_{wx}^2}$$

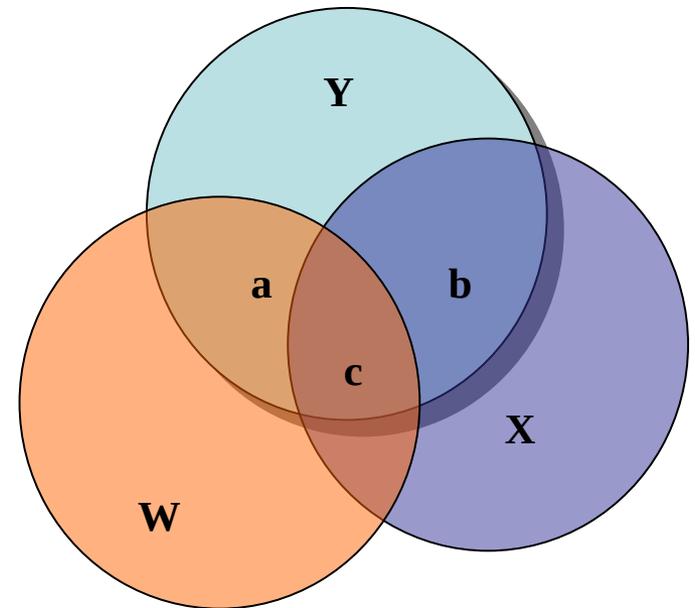


Decomposizione della varianza spiegata

- A questo punto ci possiamo chiedere quale sia l'effetto “unico” o il contributo “unico” di ogni variabile alla varianza spiegata

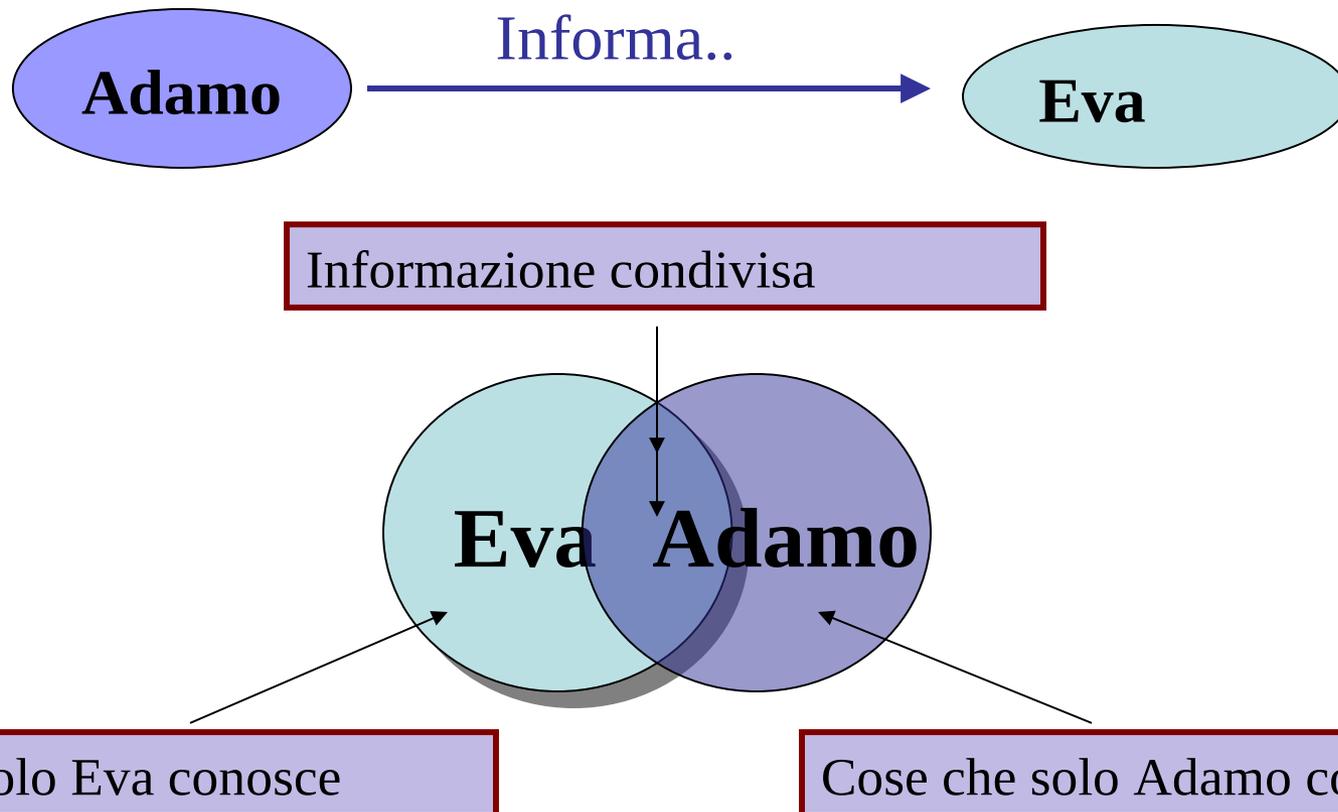
Quanto ogni VI contribuisce a spiegare varianza di Y?

Quanto è l'effetto unico di ogni variabile sulla variabilità della Y?



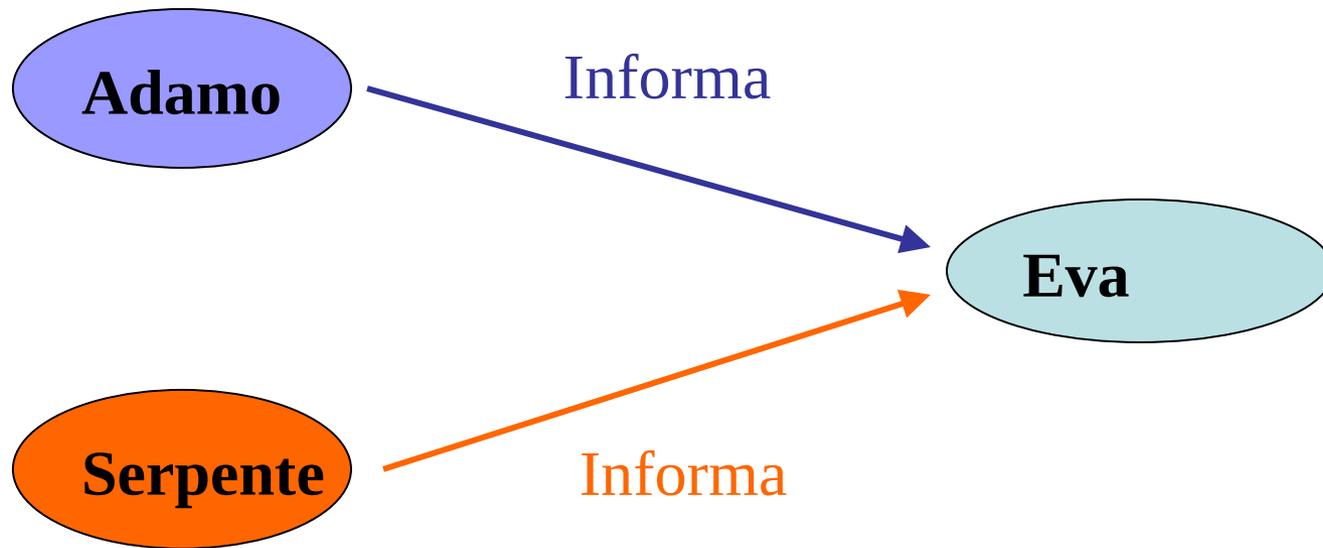
Effetti come informazione

- Immaginiamo l'effetto di una VI sulla VD come informazione trasferita dalla VI alla VD



Effetti come informazione

- Inseriamo un secondo informatore



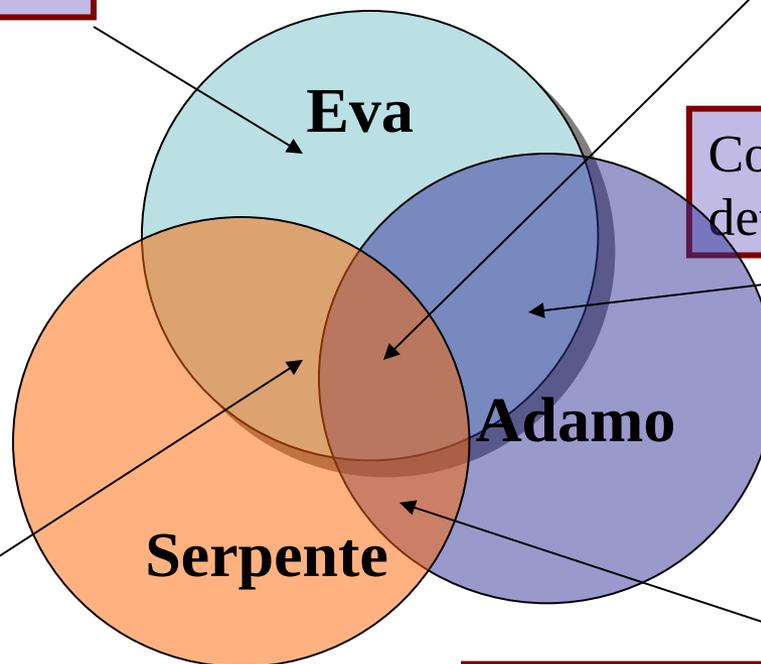
Intendiamo sapere chi ha informato Eva e
quanta parte della informazione proviene dai
due informatori

Decomposizione della informazione

- Come si ripartisce l'informazione di Eva

Cose che solo Eva
conosce

Cose che conoscono tutti



Cose che Adamo ha
detto a EVA

Cose che Serpente a detto a Eva

Cose che Adamo e Serpente
conoscono ma non Eva

Effetti unici o parziali

- Quale e' il contributo unico di Serpente alla conoscenza ottenuta da Eva?

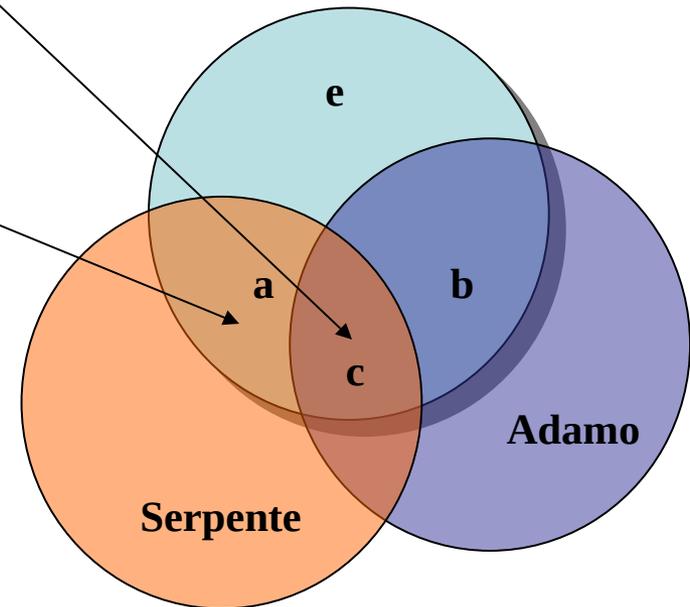
Cose che Serpente ha detto a Eva

-

Cose che tutti e tre conoscono

$$[a + c] - c = a$$

L'effetto unico di Serpente su Eva e'
dato dall'effetto totale di Serpente
rimuovendo l'effetto condiviso con
Adamo



Effetti unici o parziali

- Quale e' il contributo unico di Serpente alla conoscenza ottenuta da Eva?

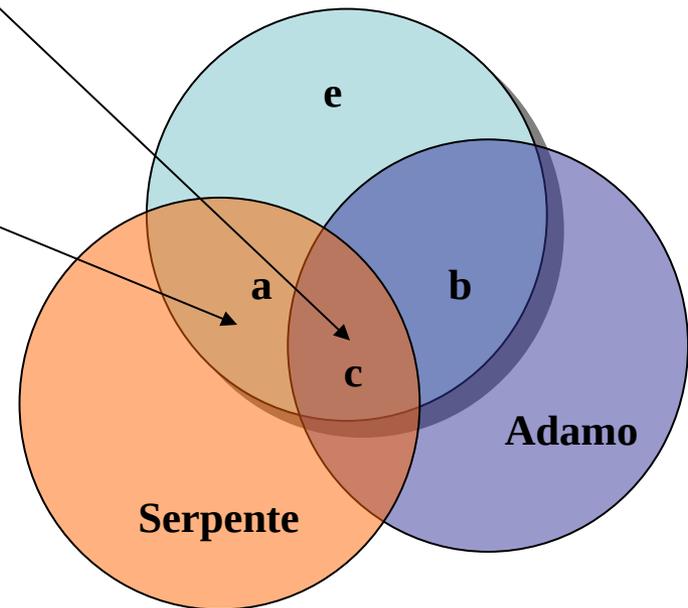
Cose che Serpente ha detto a Eva

-

Cose che tutti e tre conoscono

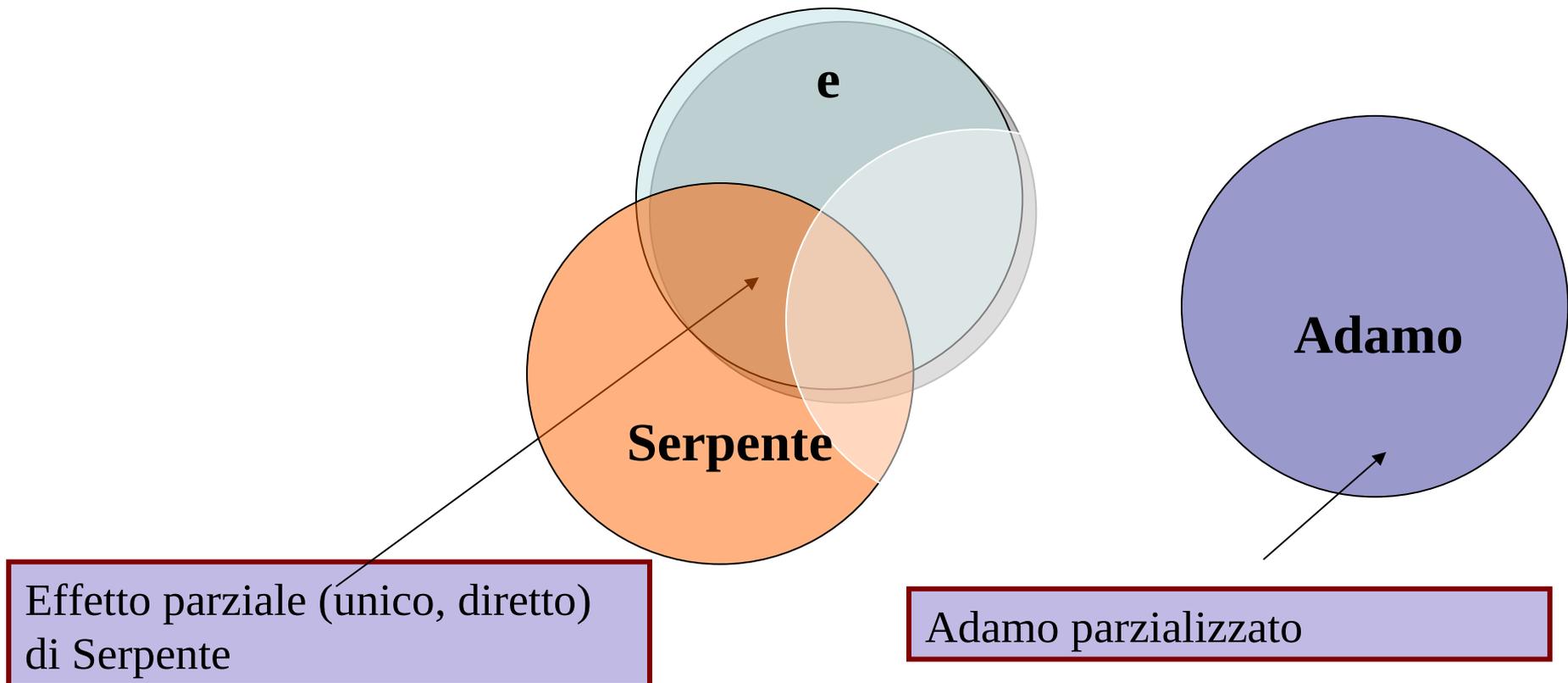
$$[a + c] - c = a$$

L'effetto unico di Serpente su Eva e'
dato dall'effetto totale di Serpente
parzializzando l'effetto condiviso con
Adamo



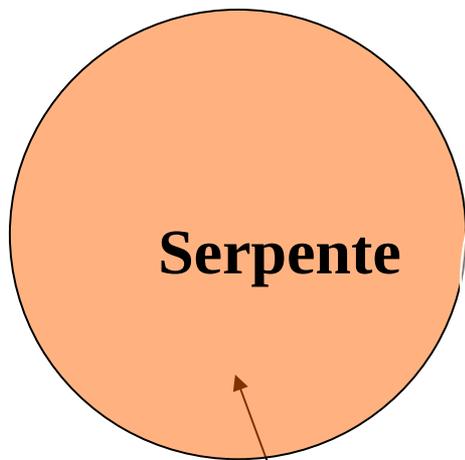
Parzializzazione

- **Parzializzare** significa rimuovere l'effetto di una (o più) VI, cioè calcolare gli effetti come se quella variabile abbia nessuna variabilità

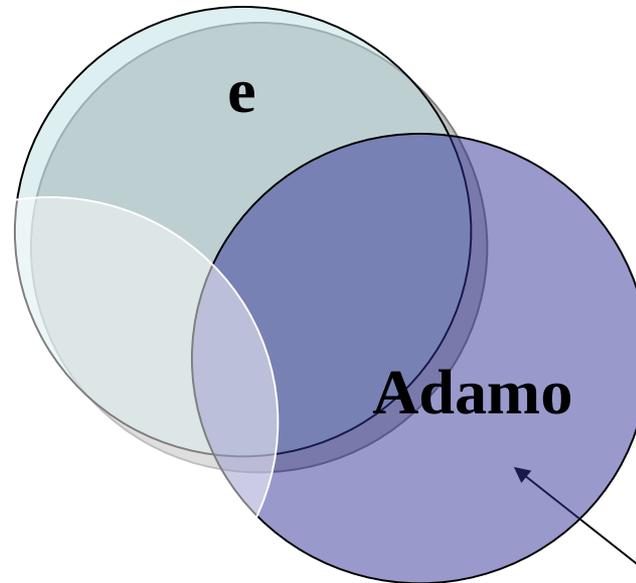


Parzializzazione

- **Parzializzare** significa rimuovere l'effetto di una (o più) VI, cioè calcolare gli effetti come se quella variabile abbia nessuna variabilità



Serpente parzializzato



Effetto parziale (unico, diretto)
di Adamo

Effetto Congiunto

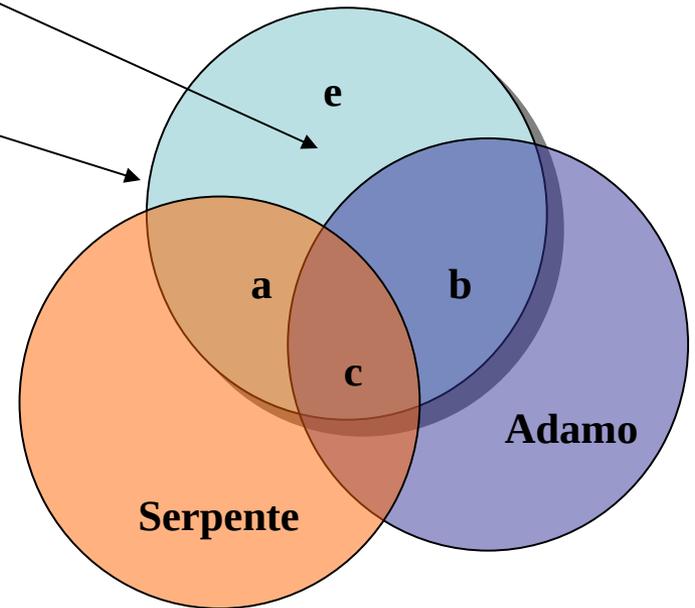
- Qual e' l'effetto totale che i due informatori hanno su Eva? Quale % della conoscenza di Eva e' dovuta agli informatori?

Conoscenza di Eva

-

Conoscenza unica di Eva

$$1 - e = a + b + c$$



Effetto Congiunto

- Qual e' l'effetto totale che i due informatori hanno su Eva? Quale % della conoscenza di Eva e' dovuta agli informatori?

Cose solo Serpente ha detto a Eva

+

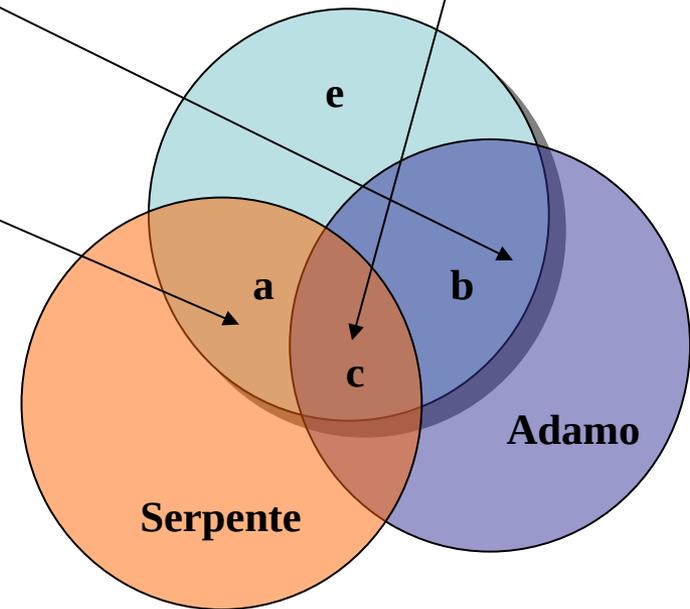
Cose solo Adamo ha detto a Eva

+

Cose che tutti sanno

$$a + b + c$$

L'effetto combinato di A e S e' dato dai loro effetti unici (parziali) e dal loro contributo comune



- Siamo interessati agli effetti di due (o piu') variabili indipendenti su una variabile dipendente
- Siamo in grado di stimare la regressione che li lega

$$Y = a + b_{yx.w} x + b_{yw.x} w + e$$

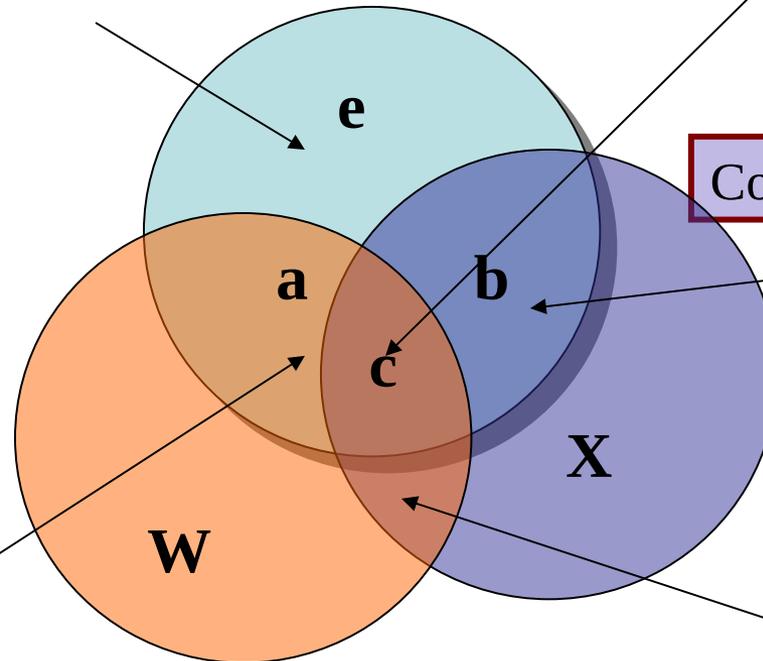
Vogliamo quantificare il loro contributo unico e combinato mediante degli indici che rappresentino la grandezza dei vari effetti

Varianza Decomposta

- Decomponiamo la varianza della variabile dipendente

Varianza di errore

Varianza completamente
condivisa



Contributo unico di X

Contributo unico di W

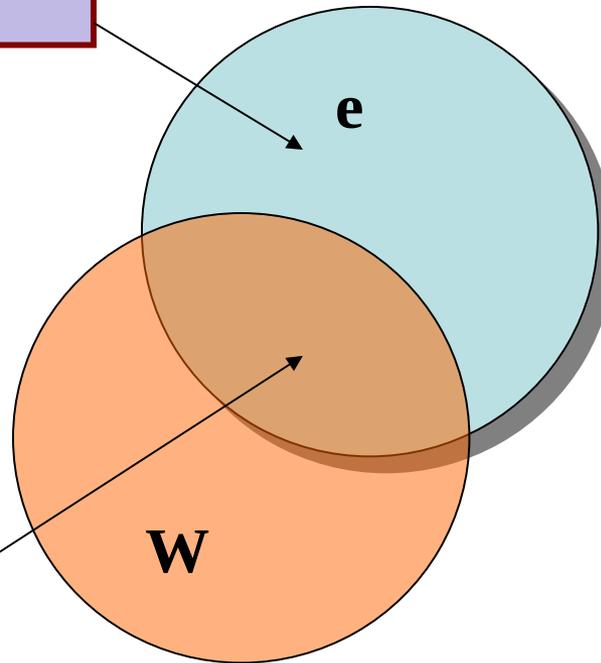
Varianza condivisa tra X e W

Correlazione semplice (quadrata)

- Ricordiamo che nella regressione semplice, la correlazione quadrata semplice e' la varianza condivisa (spiegata dalla VI)

Varianza di errore

$$r_{yw}^2$$



Contributo W

Correlazione semplice (quadrata)

- La correlazione semplice

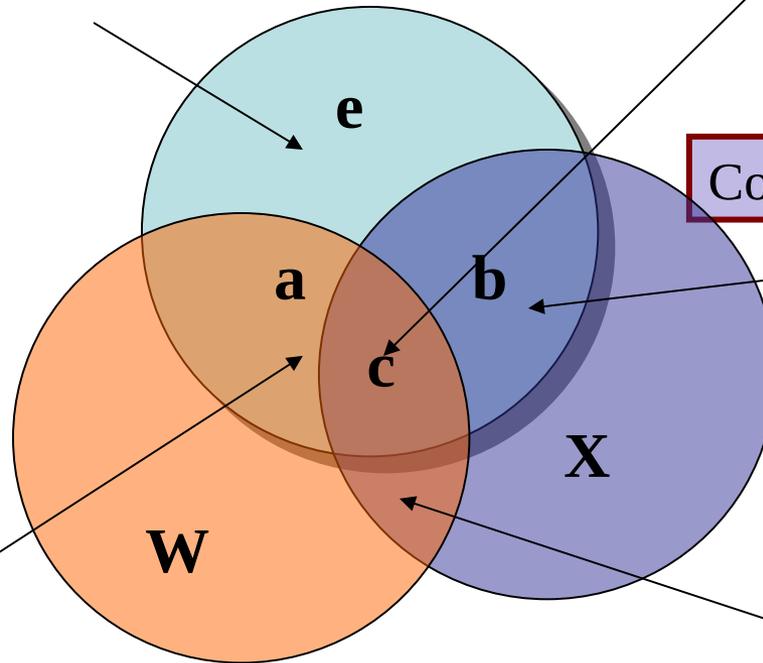
Varianza di errore

Varianza completamente condivisa

$$r_{yw}^2 = a + c$$

$$r_{yx}^2 = b + c$$

Contributo unico di X



Contributo unico di W

Varianza condivisa tra X e W

Contributo unico di VI

- Il contributo unico di una VI può essere stimato grazie al quadrato della correlazione parziale

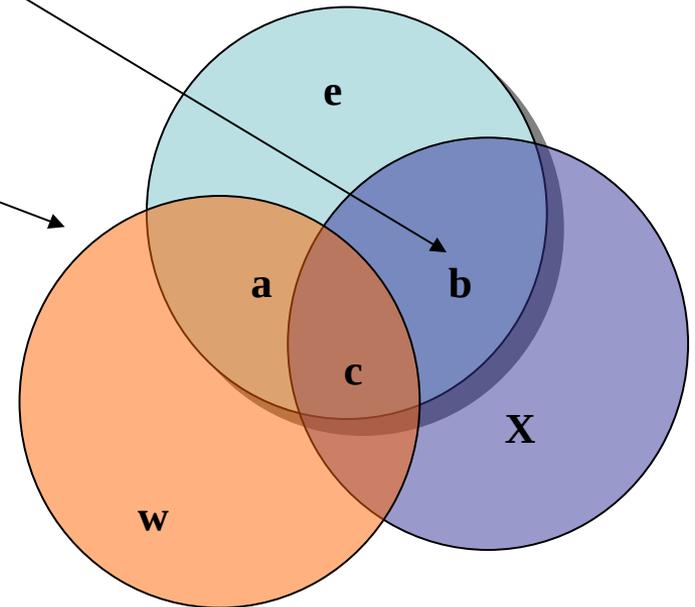
Correlazione parziale

Varianza spiegata

-

Varianza spiegata da x

$$pr_{yw.x}^2 = \frac{a}{a+e}$$



Correlazione parziale

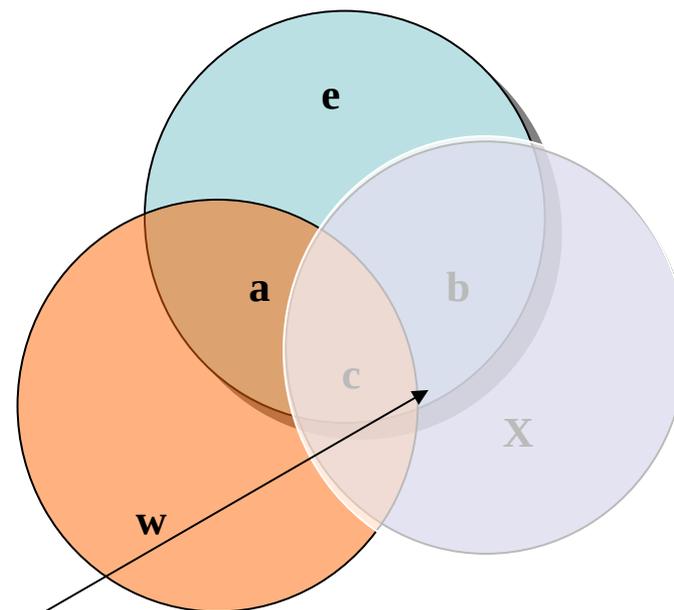
- Il quadrato della correlazione parziale indica l'effetto di una VI dopo aver rimosso tutta la variabilità delle altre

Correlazione parziale

$$pr_{yw.x}^2 = \frac{a}{a+e}$$

Varianza dovuta a w
calcolata sul totale dopo
aver tolto la varianza di x

Varianza di x e' completamente
rimossa



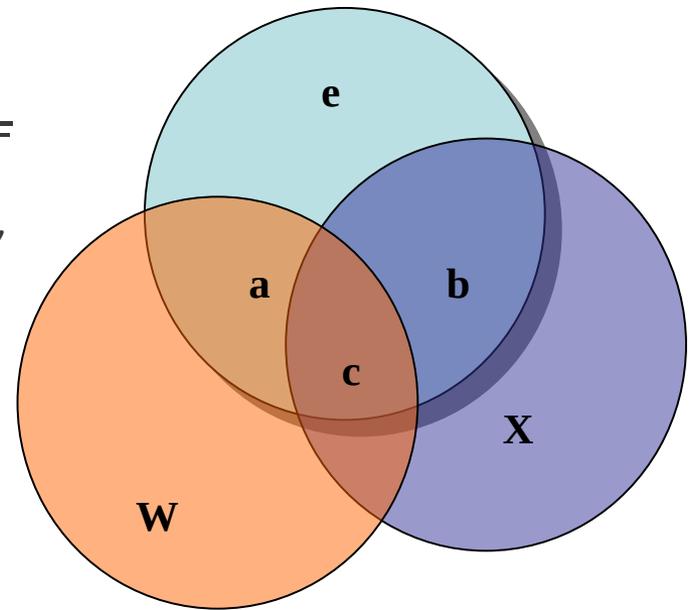
Calcolo di pr^2

- Pr può essere calcolato partendo dalle correlazioni semplici

$$pr_{yw.x} = \frac{r_{yw} - r_{yx} r_{wx}}{\sqrt{1 - r_{yx}^2} \sqrt{1 - r_{xw}^2}}$$

- Oppure dai coefficienti standardizzati

$$pr_{yw.x} = \frac{\beta_{yw.x}}{\sqrt{1 - r_{yx}^2}}$$



$$pr_{yw.x}^2 = pr_{yw.x} * pr_{yw.x}$$

Contributo unico di VI (2)

- Il contributo unico di una VI può essere anche stimato grazie al quadrato della correlazione semi-parziale

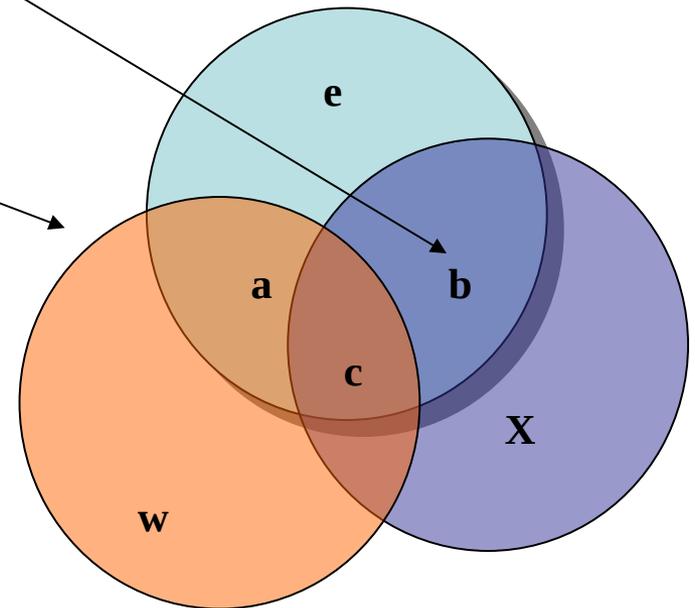
Correlazione semi-parziale

Varianza spiegata

-

Varianza spiegata da x

$$sr_{yw.x}^2 = \frac{a}{a+b+c+e}$$



Correlazione semi-parziale

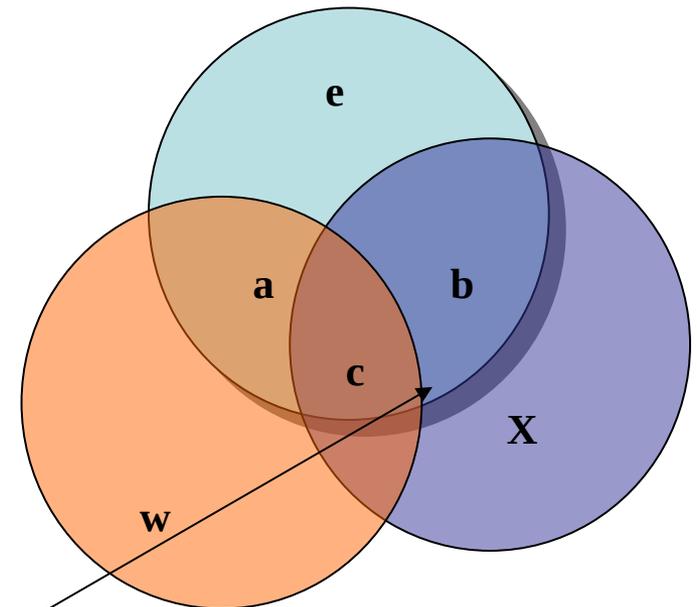
- Il quadrato della correlazione parziale indica l'effetto di una VI dopo aver rimosso tutta la variabilità delle altre

Correlazione parziale

$$sr_{yw.x}^2 = \frac{a}{a+b+c+e}$$

Varianza unica dovuta a w come proporzione di tutta la varianza di Y

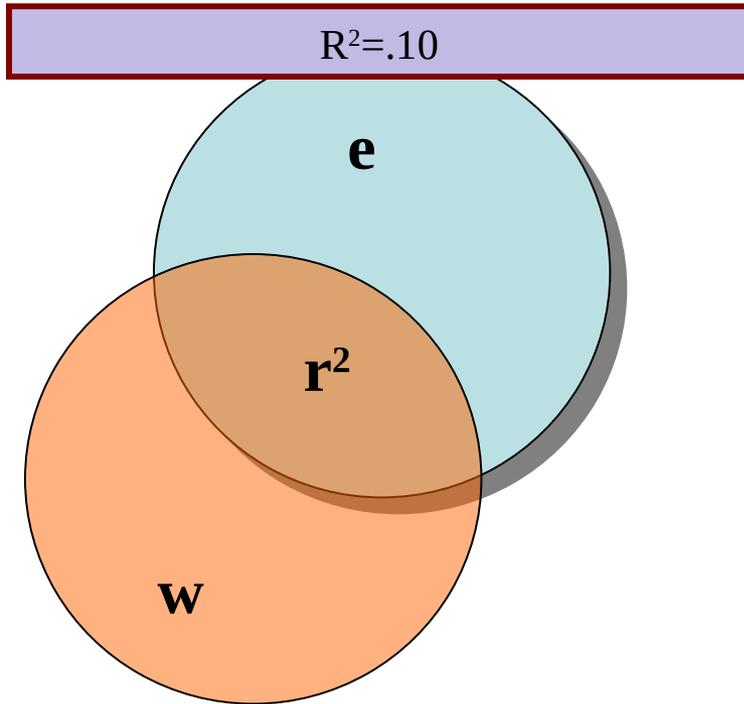
Varianza di x non rimossa nel calcolo della varianza totale



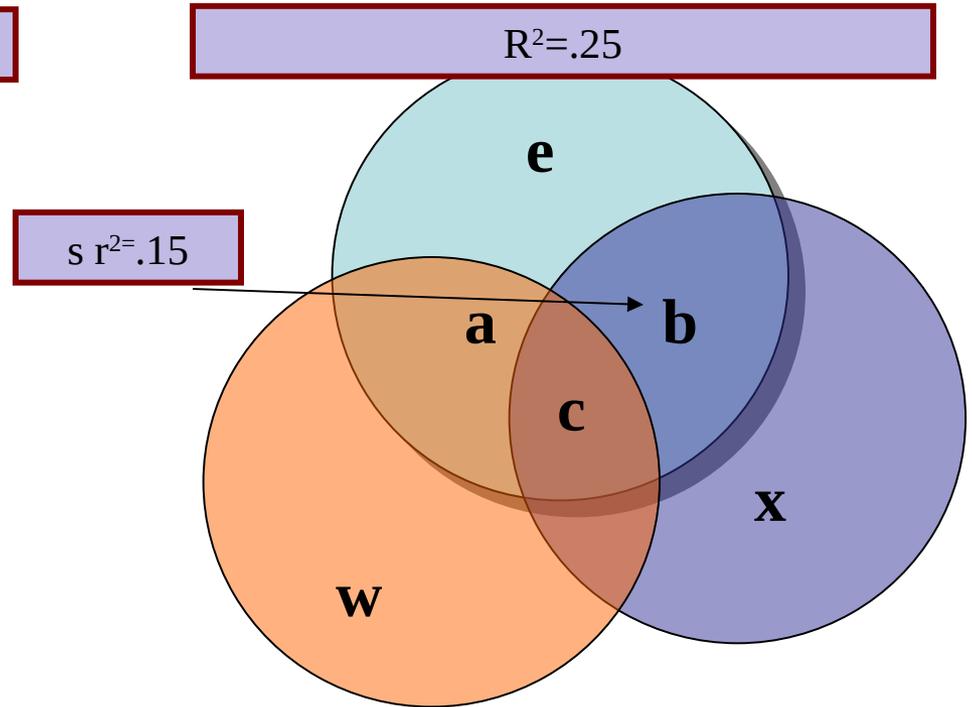
Correlazione semi-parziale

- In altri termini..

$$R_{yxw}^2 = r_{yw}^2 + sr_{yx.w}^2$$



$$.25 = .10 + .15$$



Grazie a X, si spiega un 15% in piu'

Esempio

◆ Abbiamo rilevato su una classe di 30 studenti universitari il voto preso all'esame di statistica e due variabili che vogliamo usare come predittori del voto: il numero di *ore di studio* per preparare l'esame e il numero di esami sostenuti nell'anno (*appelli*).

Statistiche descrittive

	N	Minimo	Massimo	Media	Deviazione std.
voto	30	12	30	22.17	4.942
appelli	30	0	11	4.67	2.631
ore	30	0	90	38.87	20.714
Numero di casi validi (listwise)	30				

Esempio

◆ Vediamo anche l'interazione con il software

The image shows a screenshot of a software application's menu system. The 'Analizza' menu is open, and the 'Regressione' option is highlighted. The 'Regressione' submenu is also open, showing various regression models. The 'Lineare...' option is highlighted in the submenu.

sci	Formato	Analizza	Grafici	Programmi di utilità	Estensioni	Finestra	Guida
		Report					
		Statistiche descrittive					
		Tab \grave{e} lle					
		Confronta medie					
		Modello lineare generale					
		Modelli lineari generalizzati					
		Modelli misti					
		Correlazione					
		Regressione					
		Loglineare					
		Reti neurali					
		Classifica					
		Riduzione delle dimensioni...					
		Scala					
		Test non parametrici					
		Previsioni					
		Sopravvivenza					
		Risposta multipla					
		Analisi valori mancanti...					
		Assegnazione multipla					
		Campioni complessi					
		Simulazione...					
		Controllo qualit \grave{a}					

immesse/

messe Varia

nte: voto

sse tutte le v

Riepilogo

adrato

.327

nte), ore

Somma dei quadrati

231

476

708

Modellazione lineare automatica...

Lineare...

Stima di curve...

Minimi quadrati parziali...

Logistica binaria...

Logistica multinomiale...

Ordinale...

Probit...

Non lineare...

Stima del peso...

Minimi quadrati a 2 stadi...

Scaling ottimale (CATREG)...

Esempio

- ◆ Inseriamo le variabili di interesse al posto opportuno

Regressione lineare

id
appelli
ore

Dipendente:
voto

Blocco 1 di 1

Indietro Avanti

Blocco 1 di 1

ore
appelli

Metodo: Inserisci

Variabile selezione: Regola...

Etichette del caso:

Peso WLS:

OK Incolla Reimposta Annulla Guida

Statistiche...
Grafici...
Salva...
Opzioni...
Stile...

Indipendenti

Indipendenti

Esempio

- ◆ Inseriamo le variabili di interesse al posto opportuno

Chiediamo "Statistiche"

Regressione lineare

id
appelli
ore

Dipendente:
voto

Blocco 1 di 1

Indietro Avanti

Blocco 1 di 1

ore
appelli

Metodo: Inserisci

Variabile selezione: Regola...

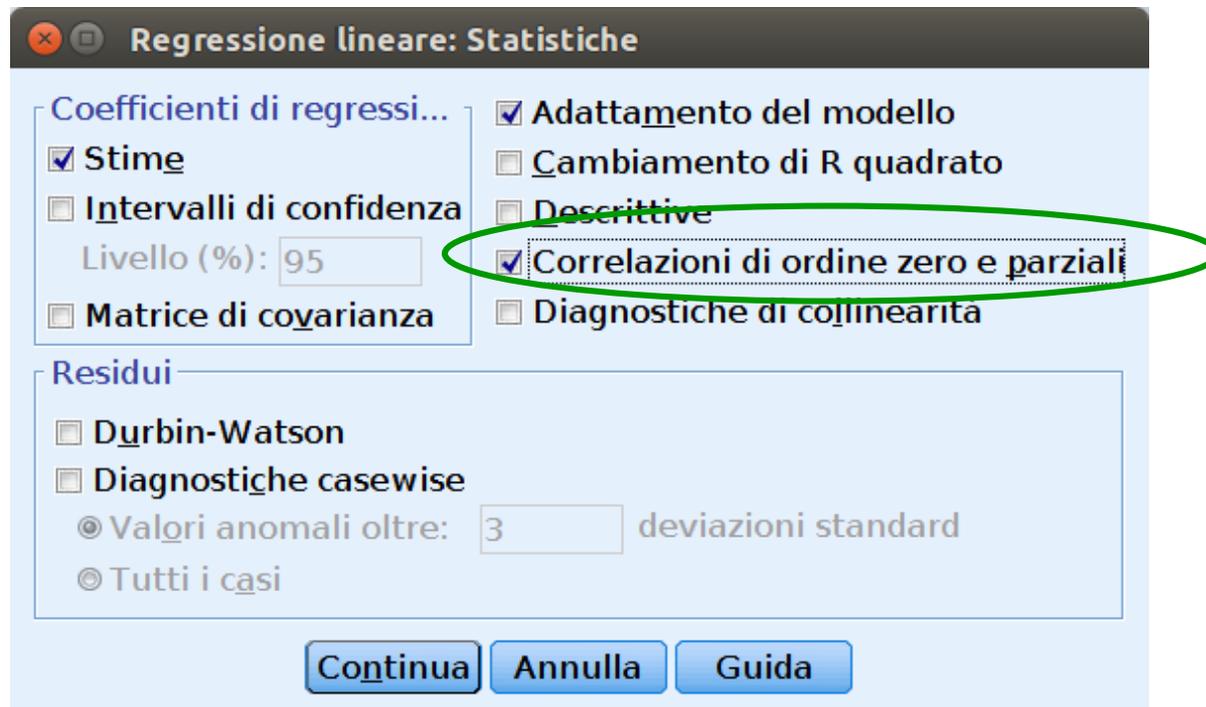
Etichette del caso:

Peso WLS:

OK Incolla Reimposta Annulla Guida

Esempio

- ◆ Chiediamo le correlazioni parziali e simi-parziali



Esempio

◆ Risultati

◆ Dobbiamo elevare al quadrato le pr e sr

Modello	Coefficienti non B					Correlazioni			
		Errore std.	Beta	t	Sign.	Ordine zero	Parziale	Parte	
1	(Costante)	18.542	2.923		6.343	.000			
	ore	.122	.043	.510	2.820	.009	.572	.477	.441
	appelli	-.236	.339	-.125	-.694	.494	-.380	-.132	-.109

	Parziale	Semi-parziale
ore	.477*.477= .227	.441*.441= .194
appelli	-.132*-.132= .017	-.109*-.109= .011

a. Variabile dipendente: voto

Part=semiparziale

Fine

Fine della Lezione III

