Elementi di Psicometria con Laboratorio di SPSS 1

12-Correlazione
(vers. 1.2, 17 novembre 2015)
versione per stampa

Germano Rossi[1]
germano.rossi@unimib.it

[1] Dipartimento di Psicologia, Università di Milano-Bicocca

2015-16

Correlazione

Correlazione di Pearson

- Capitolo 12, pp. 204-223
- escludere “La relazione lineare tra due variabili (pp. 207-212)
- escludere “Previsione e regressione lineare” (pp. 223-233)

Correlazione di Spearman

- Capitolo 21, pp. 454-457

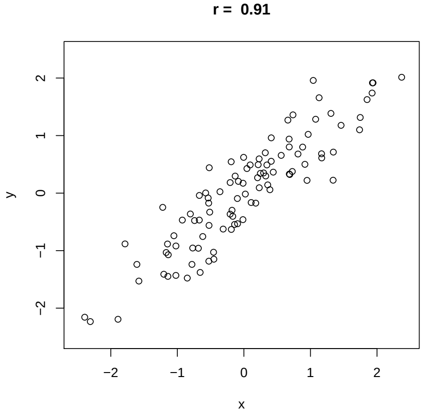
Correlazione

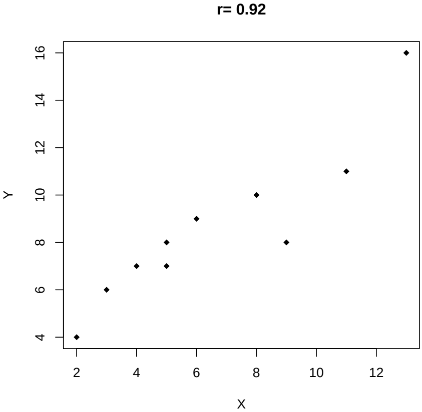
- È un indice statistico che misura l'associazione (relazione) fra due variabili
- Misura come le due variabili si muovono assieme, ossia come co-relano.
- Viene espresso come un valore che oscilla fra -1 e 1
- Per ora vedremo la correlazione lineare prodotto-momento di Bravais-Pearson più conosciuta come correlazione di Pearson e a cui ci si riferisce per antonomasia quando si parla di correlazione
- La correlazione di Pearson si usa a livello intervallo/rapporto
- Poi vedremo la correlazione di Spearman a livello ordinale (con molte categorie)

Coefficiente di correlazione

- Riassunto numerico della forza della relazione fra due variabili
- Permette di sostituire un diagramma a dispersione con un semplice indice
- È costituito da due parti:
  - Un segno che indica la direzione della relazione
  - Un numero Fra 0.00 e 1.00 che indica la forza della relazione
- 1.00 indica una relazione perfetta, esprimibile tramite una formula matematica precisa
- 0.00 indica la mancanza di qualunque relazione fra le due variabili

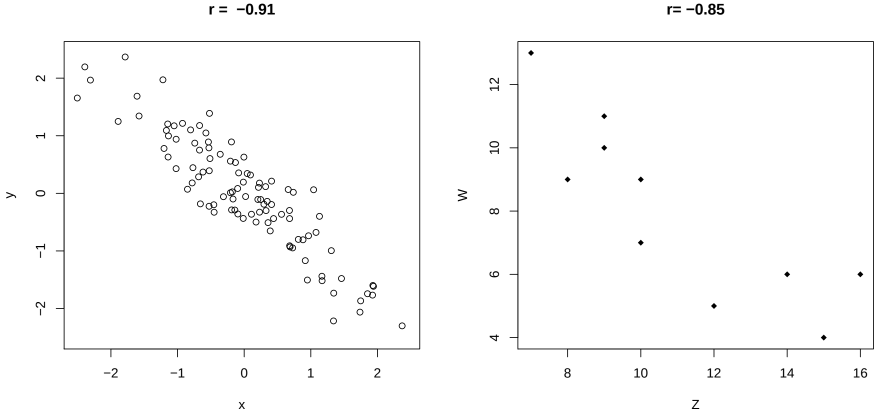
Es. di correlazione positiva





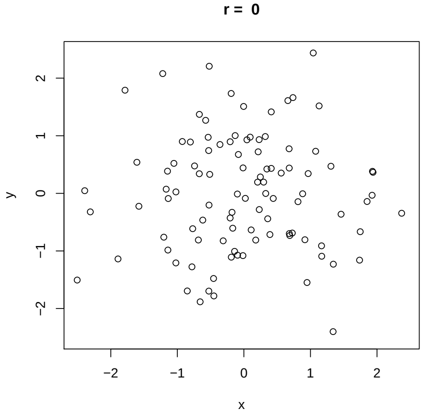
All'aumentare di X aumenta anche Y, ciascuna variabile a modo suo. E viceversa. È una relazione lineare proporzionale.

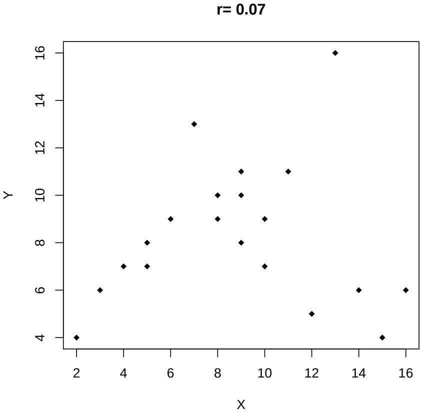
Es. di correlazione negativa



All'aumentare di X diminuisce Y, ciascuna variabile a modo suo. E viceversa. È una relazione lineare inversamente proporzionale.

Es. correlazione nulla





Non c'è alcun legame lineare fra X e Y. Ciascuna varia indipendentemente dall'altra (linearmente parlando).

Correlazione e covarianza

- Il coefficiente di correlazione è un indice che esprime la quantità di co-varianza dei dati
- rispetto al grafico a dispersione, è un indice di quanto i dati sono dispersi attorno ad una ipotetica retta che venga sovrapposta al grafico

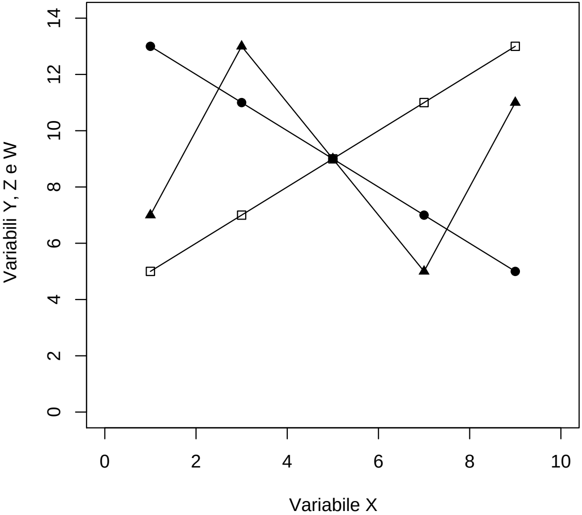
- la covarianza è un indice che esprime la quantità di varianza che due variabili anno in comune
- la formula deriva da quella della varianza

Esempio numerico

	X	Y	Z	W
a	1	5	13	7
b	3	7	11	13
c	5	9	9	9
d	7	11	7	5
e	9	13	5	11
M	5	9	9	9
s	2.828	2.828	2.828	2.828

- La relazione fra X e Y è lineare crescente ($Y = X + 4$)
- La relazione fra X e Z è lineare decrescente ($Z = -X + 14$ ovvero $Z = 14 - X$)
- La relazione fra X e W non è riconducibile ad una regola lineare

Grafico relativo



Usando gli scarti dalla media

$\overline{X} = 5$		$\overline{Y} = \overline{Z} = \overline{W} = 9$					
X		**Y**		**Z**		**W**	
1	-4	5	-4	13	4	7	-2
3	-2	7	-2	11	2	13	4
4	0	9	0	9	0	9	0
5	2	11	2	7	-2	5	-4
9	4	13	4	5	-4	11	2

k=Y, Z, W

$$(X - \overline{X})(k - \overline{k}) \quad (1)$$

$$\sum (X - \overline{X})(k - \overline{k}) \quad (2)$$

$$\frac{\sum (X - \overline{X})(k - \overline{k})}{N} \quad (3)$$

$(X - \overline{X})(k - \overline{k})$			
XY	**XZ**	**XW**	
16	-16	8	
4	-4	-8	
0	0	0	(1)
4	-4	-8	
16	-16	8	
40	-40	0	(2)
8	-8	0	(3)

1. Moltiplichiamo gli scarti fra loro
2. li sommiamo tutti
3. li dividiamo per la numerosità

In formula

- La covarianza è:

$$cov(X,Y) = \frac{\sum(X-\overline{X})(Y-\overline{Y})}{N}$$

- Mentre la varianza è:

$$var(X) = \frac{\sum(X-\overline{X})^2}{N} = \frac{\sum(X-\overline{X})(X-\overline{X})}{N}$$

Notate la somiglianza fra le due formule

Standardizziamo

Standardizziamo, dividendo per entrambe le dev. st.

$$r = \frac{cov(X,Y)}{s_x s_y}$$

$$2.828 * 2.828 = 8$$

	XY	**XZ**	**XW**
Cov	8	-8	0
s_x	2.828	2.828	2.828
s_y	2.828	2.828	2.828
$s_x s_y$	8	8	8
r	1	-1	0

Notate che una correlazione equivale ad una covarianza standardizzata sulla base delle variabili coinvolte

Formule della correlazione di Pearson

$$r = \frac{cov(X,Y)}{\sqrt{var(X)var(Y)}} = \frac{cov(X,Y)}{s_x s_y} = \frac{\frac{\sum xy}{N} - \bar{X}\bar{Y}}{s_x s_y}$$

$$r = \frac{\sum z_x z_y}{N}$$ È quella che si ricorda più facilmente

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sqrt{(\sum X^2 - \frac{(\sum X)^2}{N})(\sum Y^2 - \frac{(\sum Y)^2}{N})}}$$

$$r = \frac{N\sum XY - \sum X \sum Y}{\sqrt{[N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2]}}$$

Calcolo d'esempio: manuale

	X	Y	X^2	Y^2	XY
a	1	5	1	25	5
b	3	7	9	49	21
c	5	9	25	81	45
d	7	11	49	121	77
e	9	13	81	169	117
$\sum$	25	45	165	445	265

$$\frac{5 \cdot 265 - 25 \cdot 45}{\sqrt{[5 \cdot 165 - 25^2]\,[5 \cdot 445 - 45^2]}} = \frac{200}{\sqrt{200 \cdot 200}} = 1$$

Interpretazione

L'interpretazione si applica al valore della correlazione indipendentemente dal segno

Valore di r	Correlazione	Relazione
0.00-0.20	Piccola	Molto poco intensa, quasi inesistente
0.20-0.40	Bassa	Piccola, appena appena apprezzabile
0.40-0.60	Regolare	Considerevole
0.60-0.80	Alta	Intensa
0.80-1.00	Molto alta	Molto intensa

Il segno indica solo la relazione proporzionale o inversamente proporzionale

Correlazione

- Immaginate di aver raccolto un campione di 20 persone
- di aver misurato 2 variabili
- e di aver trovato un valore di .56

- In termini assoluti è una buona correlazione ma...

- Siamo sicuri che il valore di .56 con un campione di 20 persone
- sia una buona stima della correlazione della popolazione?
- Potrebbe essere un campione "balordo" con una correlazione eccessivamente alta (o bassa)

- Usiamo la logica della distribuzione campionaria

Distribuzione campionaria della correlazione

- Usiamo una popolazione finita di 2 variabili che correlano a 0.00365
- Estraiamo dei campioni di ampiezza 20
- Calcoliamo la correlazione per ciascuno dei campioni…
- Facciamo la rappresentazione grafica per vedere come:
  - i valori vicini a 0 sono i più frequenti
  - valori (positivi e negativi) vicini a 0 sono leggermente meno frequenti di 0
  - man mano i valori si allontanano da 0, meno frequenti diventano
- In pratica i valori della distribuzione campionaria della correlazione si distribuiscono approssimativamente come una normale.

Distribuzione campionaria della correlazione

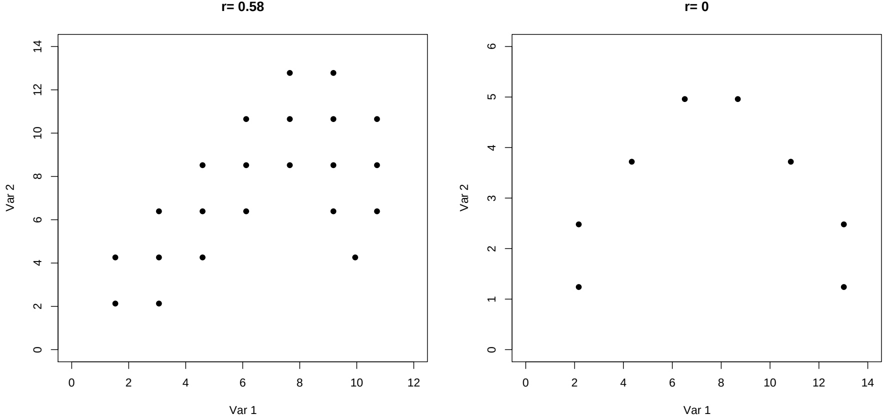
- L'approssimazione alla normale è sempre migliore all'aumentare dell'ampiezza dei campioni (per N piccole si può aggiustare la distribuzione)
- Se la correlazione trovata nel nostro campione di partenza è compresa nel 95% attorno alla media di 0, allora la nostra correlazione sarà **non significativa** ovvero casualmente estratta da una popolazione con correlazione 0
- Se la correlazione trovata sarà compresa nel 5% delle due code della normale, allora sarà considerata significativa, cioè un valore poco probabile da ottenere casualmente**.**

Dimostrazione (o verifica) dell'ipotesi

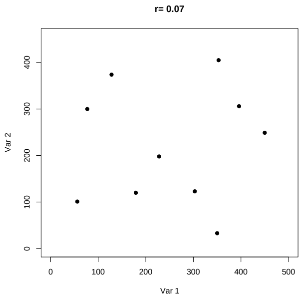
- Quello che abbiamo fatto può essere considerato una dimostrazione di un'ipotesi
- Abbiamo ipotizzato che nella popolazione da cui abbiamo estratto il campione, la correlazione fra le due variabili sia 0
- Abbiamo costruito una distribuzione campionaria della correlazione
- E abbiamo confrontato la correlazione calcolata con la distribuzione delle correlazioni
- Se la probabilità associata alla nostra correlazione è $\leq$ 2.5% allora riteniamo che sia improbabile che il nostro campione sia stato estratto da quella popolazione (che ha r=0)
- In tal caso, concludiamo che il campione viene da una popolazione diversa

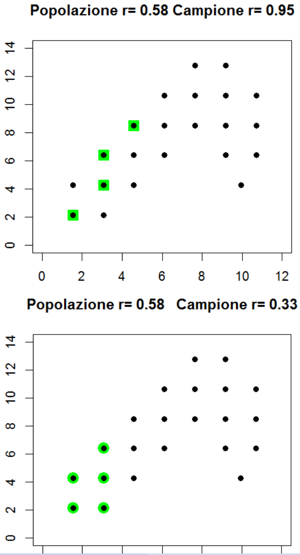
Correlazione lineare

Se i dati non sono “lineari” la correlazione non è “buona”; la relazione potrebbe non essere affatto lineare. Per questo l’inferenza sulla correlazione verifica che sia estratta da una popolazione con correlazione nulla, cioè $H_0 : \rho = 0$ (rho)



Inferenza





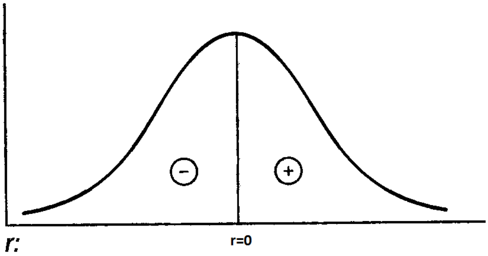
- Nel primo caso, la relazione non è lineare ma il campione che abbiamo estratto (quadrati) ce lo fa credere: r=.95
- Nel secondo (cerchi) è il contrario: r=-.08

Inferenza

Nell'inferenza per la correlazione, H_0 è sempre uguale ($H_0 : \rho = 0$), mentre le ipotesi alternative potrebbero essere:

- $H_1 : \rho \neq 0$
- $H_1 : \rho > 0$
- $H_1 : \rho < 0$

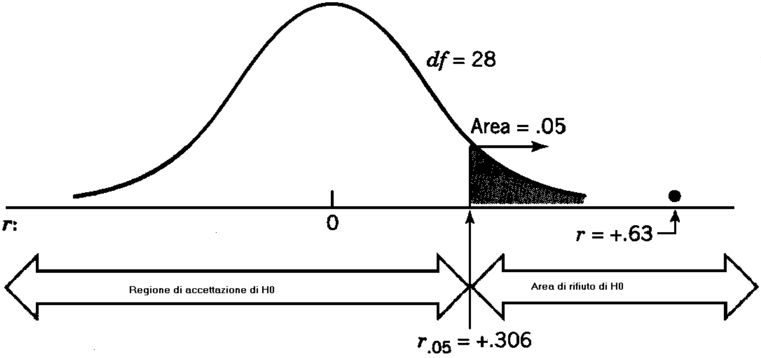
con gl $= N - 2$



In pratica ci chiediamo se il valore da noi trovato viene da una popolazione con correlazione nulla. **Se accettiamo** H_0, sì e quindi **la correlazione trovata** (qualunque sia il suo valore) **non deve neppure essere presa in considerazione** (non va interpretata).

Inferenza

Nel caso di un'ipotesi monodirezionale positiva



Inferenza: uso delle tavole

La tavola dei valori critici riporta i valori (per i gradi di libertà, per diversi α e per le due ipotesi, mono e bi-direzionali) sotto i quali accettare l'ipotesi nulla.

Valori critici del coefficiente *r* di Pearson***r***

df **(= *N* = 2; *N* = numero di coppie di dati)**	**Livello di significatività per il test a una coda**			
	0,05	**0,025**	**0,01**	**0,005**
	Livello di significatività per il test a due code			
	0,10	**0,05**	**0,02**	**0,01**
1	0,988	0,997	0,9995	0,9999
2	0,900	0,950	0,980	0,990
3	0,805	0,878	0,934	0,959
4	0,729	0,811	0,882	0,917
5	0,669	0,754	0,833	0,874

$N = 7 \qquad \alpha = .05$(bi) $\qquad r_t = .65 \Rightarrow H_0 \qquad r_t = .79 \Rightarrow H_1$

Esempio numerico

X	Y	X^2	Y^2	XY
46	126	2116	15876	5796
49	110	2401	12100	5390
48	103	2304	10609	4944
42	128	1764	16384	5376
46	111	2116	12321	5106
49	128	2401	16384	6272
43	104	1849	10816	4472
45	101	2025	10201	4545
49	111	2401	12321	5439
42	125	1764	15625	5250
40	113	1600	12769	4520
45	115	2025	13225	5175
48	100	2304	10000	4800
41	124	1681	15376	5084
43	101	1849	10201	4343
40	102	1600	10404	4080
47	129	2209	16641	6063
48	112	2304	12544	5376
48	128	2304	16384	6144
46	123	2116	15129	5658
905	2294	41133	265310	103833

Esempio numerico

$$\frac{20 \cdot 103833 - 905 \cdot 2294}{\sqrt{(20 \cdot 41133 - 905^2)(20 \cdot 265310 - 2294^2)}} =$$

$$\frac{2076660 - 2076070}{\sqrt{(822660 - 819025)(5306200 - 5262436)}} =$$

$$\frac{590}{\sqrt{3635 \cdot 43764}} = \frac{590}{\sqrt{159082140}} = \frac{590}{12612.7768524} = 0.0467$$

- $H_1 : \rho \neq 0$
- Gdl: $20 - 2 = 18$

α	.05	.01
r_c	.444	.561

Matrice varianza/covarianza

- Nel futuro lavorerete molto con le varianze e covarianze
- Spesso in una forma particolare: la tabella varianze/covarianze
- Varianze lungo la diagonale, covarianze fuori
- Idem per una tabella di correlazioni

VAR	Authori	Malleab	World
Authori	**38.694**	16.891	16.733
Malleab	16.891	**37.106**	10.137
World	16.733	10.137	**33.323**

COR	Authori	Malleab	World
Authori	**1.000**	0.446	0.466
Malleab	0.446	**1.000**	0.288
World	0.466	0.288	**1.000**

$$cov(X, X)/\sqrt{s * s} = 1$$

Correlazione in SPSS

- `Analizza | Correlazione | Bivariata`
- Dal riguadro "Coefficienti di correlazione" scegliere **Pearson** o **Spearman**
- Il riquadro "Test di significatività" permette di scegliere l'opzione bidirezione ("A due code", preferibile) o monodirezionale ("A una coda")
- Il riquadro "Evidenzia correlazioni significative" permette di aggiungere degli asterischi di significatività
- Il `Opzioni` permette di scegliere le opzioni "Esclusioni a coppie" o "Esclusione listwise" per i valori mancanti
- Lo stesso bottone permette di chiedere le statistiche descrittive (media e dev.st.) e la covarianza

Matrice correlazioni completa

		PIACERE	DOLORE	TIMORE	DIVERTIM	ALLEGRIA	IRRITAZI	MALINCON
PIACERE	Correlazione di Pearson	1						
	Sig. (2-code)	.						
	N	74						
DOLORE	Correlazione di Pearson	,077	1					
	Sig. (2-code)	,514	.					
	N	74	74					
TIMORE	Correlazione di Pearson	,048	,236(*)	1				
	Sig. (2-code)	,684	,043	.				
	N	74	74	74				
DIVERTIM	Correlazione di Pearson	,330(**)	,097	,140	1			
	Sig. (2-code)	,004	,413	,235	.			
	N	74	74	74	74			
ALLEGRIA	Correlazione di Pearson	,310(**)	,041	,119	,737(**)	1		
	Sig. (2-code)	,007	,727	,312	,000	.		
	N	74	74	74	74	74		
IRRITAZI	Correlazione di Pearson	-,401(**)	,073	,068	-,150	-,170	1	
	Sig. (2-code)	,000	,534	,563	,202	,147	.	
	N	74	74	74	74	74	74	
MALINCON	Correlazione di Pearson	,149	,415(**)	,278(*)	,094	,005	,025	1
	Sig. (2-code)	,204	,000	,016	,426	,964	,831	.

Matrice correlazioni compatta

		Religiosità				Attaccamento		
	Fondamen-talismo	intrinseca	estrinseca sociale	estrinseca personale	Ortodossia religiosa	sicuro	preoccupato	spaventato
Fondamen-talismo	1							
Religiosità intrinseca	,679(**)	1						
Religiosità estrinseca sociale	,310(**)	,343(**)	1					
Religiosità estrinseca personale	,510(**)	,523(**)	,353(**)	1				
Ortodossia religiosa	,727(**)	,704(**)	,203(**)	,620(**)	1			
Attacc. sicuro	,116(*)	,108(*)	,016	,042	,110(*)	1		
Attacc. preoccupato	,089	,052	,060	,088	,084	-,223(**)	1	
Attacc. spaventato	-,037	-,084	-,068	,062	-,027	-,195(**)	,241(**)	1
Attacc. distanziante	-,209(**)	-,258(**)	-,140(**)	-,066	-,157(**)	-,092	-,205(**)	,406(**)

Legame di `r` con `t` appaiato

- Sia il `t` appaiato che `r` utilizzano le varianze (o le dev. st)
- Esiste una formula che permette di calcolare `t` usando `r`

$$t = \frac{X_1 - X_2}{\sqrt{\frac{s_1^2 + s_2^2}{N} - \frac{2 r s_1 s_2}{N}}}$$

- Non è una formula molto utile, ma ci permette di vedere che una parte del `t` appaiato è legato a `r`
- All'aumentare di `r`, `t` aumenta
- Quanto $r = 0$, `t` diventa uguale a `t` per campioni indipendenti

Attendibilità e validità

La correlazione (o misure basate sulla correlazione) vengono usate per l'analisi delle scale che andranno a formare un test psicologico

- **Attendibilità** o **Affidabilità**: quanto ci si può fidare che lo strumento misuri fedelmente ogni volta?
    - **Test-retest**: lo stesso strumento si somministra due volte a distanza di tempo, la correlazione fra le due dev'essere almeno .70
    - **Split-half**: lo strumento viene diviso in due parti (item pari e item dispari) e i punteggi delle due metà vengono confrontati fra loro
    - **Alfa di Cronbach**: è un indice basato su tutte le correlazioni possibili fra gli item della scala. Deve essere almeno .70 (.60 con pochi item)

Attendibilità e validità

- **Validità**: stiamo misurando veramente quello che pensiamo di misurare?
  - Dipende da cosa stiamo facendo
  - **Nuova versione** di uno strumento: deve avere una correlazione elevata con la vecchia versione
  - **Nuovo strumento** per un costrutto mai misurato prima: deve correlare abbastanza/molto con altre misure che si ipotizzano siano correlate al costrutto

Attendibilità in SPSS

`Analizza | Scala | Analisi di affidabilità`

Split-half

	Split-half	N di item
Scala di ortodossia RFS	,740	9

Alfa di Cronbach

	Alfa di Cronbach	N di item
Scala di ortodossia RFS	,885	9

Per entrambe le procedure, gli item contro-tratto devono essere ribaltati

Validità in SPSS

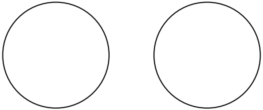
Validità con criterio esterno

		Scala di fondamentalismo RF
Scala di ortodossia RFS	r	,727
	Sig.	,000

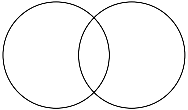
Il fondamentalismo include l'ortodossia

Rappresentazione grafica di r

- due variabili NON correlate



- due variabili correlate



- L'area in comune rappresenta la varianza che le due variabili condividono fra loro
- In termini di contenuto è qualcosa che è misurato contemporaneamente da entrambe le variabili

Coefficiente di determinazione

- La correlazione indica quanto sono associate le variabili
- Il quadrato della correlazione indica esattamente quanta varianza hanno in comune le variabili
- Se poi si moltiplica per 100 si ha la % di varianza comune

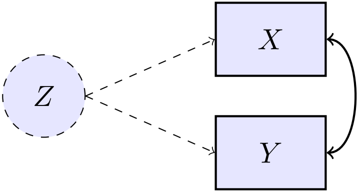
$$r = .9 \quad r^2 = .81, 81\% \qquad r = .7 \quad r^2 = .49, 49\%$$

$$r = .6 \quad r^2 = .36, 36\% \qquad r = .4 \quad r^2 = .16, 16\%$$

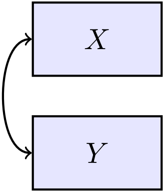
$$r = .3 \quad r^2 = .09,\ 9\% \qquad r = .2 \quad r^2 = .04,\ 4\%$$

Legame fra le variabili

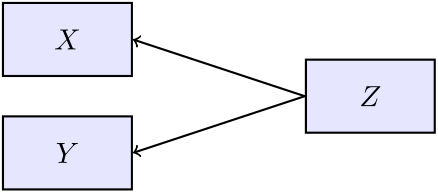
- È importante ricordare che se esiste una correlazione fra due variabili, che calcoliamo con r, questo indice non ci dà nessuna informazione sui legami di causa-effetto.
- Le due variabili "si muovono assieme". STOP!
- È possibile che esista una terza variabile che ha influenza su entrambe e che la correlazione che abbiamo calcolato sia dovuta a questa influenza



False correlazioni



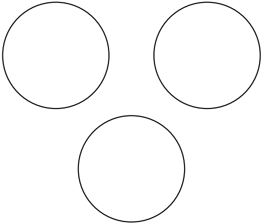
- È falsa una correlazione esistente che non ha senso logico ma che può portare ad una interpretazione apparentemente "accettabile"
- X è il numero di pompieri mandato a spegnere un incendio
- Y è l'entità del danno prodotto dall'incendio
- La loro correlazione vuol dire che più pompieri producono più danni?



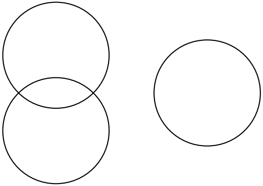
- Nel momento in cui si identifica una variabile antecedente ad entrambe, la correlazione spuria acquista senso
- Z è l'ampiezza dell'incendio
- Più ampio l'incendio, più pompieri vengono inviati a spegnerlo
- più ampio l'incendio, più danni prodotti

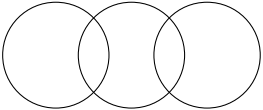
Con tre variabili

$r = 0$

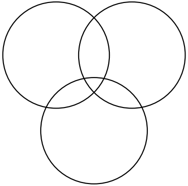


2 correlano fra loro





1 correla con le altre due



correlano tutte

Correlazione multipla

- È la correlazione di una variabile con 2 o più variabili contemporaneamente
- Oscilla fra -1 e $+1$ come la correlazione di Pearson (come tutti gli indici di correlazione)

$$r_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12}r_{13}r_{23}}{1 - r_{23}^2}}$$

dove r_{12} è la correlazione fra le variabili 1 e 2; r_{13} fra la 1 e la 3… e $r_{1.23}$ è la correlazione multipla

In SPSS si ottiene solo come sottoprodotto della regressione lineare multipla

Correlazione parziale

- È la correlazione di due variabili a cui viene “tolta” l’influenza di una terza variabile. In pratica si cerca di scorporare l’influenza di terza, quarta… variabile per trovare la relazione “vera”
- Es. correlazione fra “numero di parole conosciute” da un bambino e “intelligenza” parzializzata in base all’età (tolto il contributo dell’età). Se l’età è correlata con una delle due o con entrambe, la correlazione diminuirà.

$$r_{12.3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{(1 - r_{13}^2)(1 - r_{23}^2)}}$$

dove r_{12} è la correlazione fra le variabili 1 e 2; r_{13} fra la 1 e la 3… e $r_{12.3}$ è la correlazione fra 1 e 2 parzializzata sulla 3

Correlazione semi-parziale

- È la correlazione fra due variabili, ma solo ad una delle due è stato tolto il contributo di una terza.
- Es. correlazione fra "numero di parole conosciute" e "intelligenza". La parzializzazione in base all'età viene attuata solo con il numero di parole.

$$r_{1(2.3)} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1 - r_{23}^2}}$$

dove r_{12} è la correlazione fra le variabile 1 e 2, r_{13} fra la prima e la terza e così via

In SPSS non è possibile ottenere la correlazione semi-parziale, se non come risultato opzionale di una regressione multipla

Correlazione parziale in SPSS

- `Analizza | Correlazione | Parziale`
- Nel riquadro “Variabile” inserire almeno due nomi di variabile
- Nel riquadro “Rimuivi effetti di” inserire le variabili di cui vuole eliminare l’effetto
- Nell riquadro “Test di significatività” scegliere l’opzione bidirezione (“A due code”, preferibile) o monodirezionale (“A una coda”)
- Il riquadro “Mostra l’esatto livello di significativà” passa dalla modalità compatta (asterischi di significatività) a quella completa (se abilitato)
- Il `Opzioni` permette di scegliere le opzioni “Esclusioni a coppie” o “Esclusione listwise” per i valori mancanti
- Lo stesso bottone permette di chiedere le statistiche descrittive (media e dev.st.) e la correlazione di ordine zero

	Correlazioni con Fondamentalismo	
	Ordine zero	Parziali
Orient. Polit.	0.281	0.107
Rel. Intrinseca	0.679	0.274
Rel. estr. pers.	0.310	0.145
Rel. Estr. soc	0.510	0.026
Ortodossia	0.727	0.422
Attac. Sicuro	0.115	0.039
Attac. Preoccup.	0.089	0.024
Attac. Spavent.	-0.037	0.035
Attac. Distanz.	-0.209	-0.063

La correlazione di ordine zero è la normale correlazione
La correlazione parziale (in questo caso) è parzializzata su tutte le altre

Correlazione di Spearman

- Si usa con variabili ordinali o rese ordinali (ad es. variabili intervallo discrete con N piccolo o andamento “non normale”, ad es. campioni patologici)

X	Y
A	3
B	3
A	1
D	2
C	3
B	2

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

Trasformazione in ranghi

valori	A	A	B	B	C	D
pos.	1	2	3	4	5	6
rango	1.5	1.5	3.5	3.5	5	6
valori	1	2	2	3	3	3
pos.	1	2	3	4	5	6
rango	1	2.5	2.5	5	5	5

A valore uguali deve essere assegnato lo stesso rango, quindi ai due valori A viene assegnato la media dei due ranghi $(1+2)/2 = 1.5$

Correlazione di Spearman

- Ordinare i valori
- Indicare la posizione
- Assegnare i ranghi
- La somma dei ranghi deve coincidere

	X	rango X	Y	rango Y	d	d^2
	A	1.5	3	5	-4	12
	B	3.5	3	5	-2	2,3
	A	1.5	1	1	0,5	0,3
	D	5	2	2.5	2,5	6,3
	C	6	3	5	1	1
	B	3.5	2	2.5	1	1
$\sum$		21		21		23

$$r_s = 1 - \frac{6 \cdot 23}{6(6^2-1)} = 1 - \frac{138}{6 \cdot 35} = 1 - \frac{138}{210} = 1 - 0.657143 = 0.343$$

Altre misure di correlazione

Coefficiente di correlazione	**Livelli di misurazione**
Prodotto momento di *Pearson*	Entrambe intervallo
A ranghi di *Spearman*, Tau di *Kendall*	Entrambe ordinali
Phi, *V di Cramer*	Entrambe nominali
Punto-biseriale	Una intervallo e una dicotomica vera
Biseriale*	Una intervallo e una dicotomica artificiale
Contingenza	Entrambe nominali
Tetracorica*	Entrambe dicotomiche artificiali
Poliseriale	Una intervallo e una ordinale
Policorica*	Entrambe ordinali artificiali

In corsivo quelle ottenibili con SPSS
L'asterisco indica quelle ottenibili con macro SPSS scaricabili da internet